

# Theological Determinism and the Goodness of God

A thesis submitted for the degree of Doctor of Philosophy

by

Matthew James Hart

University of Liverpool

September 2019

## Abstract

This thesis is a defence of theological determinism (TD) as it is expressed in a Reformed or Calvinistic context. TD is the doctrine that God determines everything that occurs. It is not a popular view amongst theistic philosophers at the present time—most theists are *libertarians*. That is to say, most contemporary theistic philosophers think that moral responsibility is incompatible with determinism (they are *incompatibilists*), and that man is a morally responsible creature. Theological determinists, on the other hand, are typically *compatibilists*. They agree that man is a morally responsible creature, but they think this claim is consistent with determinism.

My task in this thesis is to defend TD against the two signal classes of objection that have been brought against it: (i) arguments for incompatibilism that urge that TD removes human moral responsibility, and (ii) arguments to the effect that that TD calls into question the goodness, love and paternity of God on account of the closer connection between God and evil that TD posits.

The structure of the thesis is as follows.

In Part 1 I set the stage. In chapter 1 the different proposed analyses of determinism are laid out, and I offer my preferred understanding. I suggest that the standard definitions of determinism are inadequate. The crucial notion is not entailment, as the standard definitions suppose, but some sort of metaphysically prior, explanatory determination. The various terms of the free-will debate (compatibilism, incompatibilism, libertarianism, etc.) are explicated in chapter 2.

Chapter 3 is given over to a historical centring of the debate. I don't want to discuss theological determinism in the abstract, but in the light of the Reformed or Calvinist tradition. The two opposing schools of thought in this regard are Calvinism and Arminianism. I survey the history of the debate starting with Luther and Erasmus's dispute over free will, and trace the formation of the two opposing points of view all the way up to the *Westminster Confession*. My contention is that Calvinism should be understood as a compatibilistic and deterministic viewpoint, in contrast to the incompatibilistic and indeterministic system of Arminianism. The great objections the Arminians have traditionally made are also presented.

In Part 2 I begin to respond to those objections. Part 2 is given over to arguments against TD that are premised on arguments against compatibilism. Chapter 4 is my response to the Ability Argument against determinism, which has it that the ability to do otherwise is necessary for moral responsibility and that determinism removes that ability. I respond by suggesting that the relevant 'could have done otherwise' language might refer to compossibility, not ability, and the relevant compossibility is simply that of one's desire-set with a different decision (intention). Thus, the

Calvinist may carry on assuming ‘could have done otherwise’ language is relevant to moral responsibility.

Chapter 5 concerns the Manipulation Argument. This argument is a recent development, and it appeals to the intuition that if an individual causes you to do something (even if through a long and distant chain of causation), then you are not responsible for what you were thus caused to do. But TD therefore implies that no human being is morally responsible, because God is the ultimate cause of all our actions. I respond that manipulation cases don’t show us an absence of responsibility, they show us an uncertainty on the given facts, but that uncertainty can be embraced by the Calvinist, for he can appeal to divine testimony, a further fact, to justify his belief that human beings are morally responsible. I also turn the tables and offer my own manipulation argument for compatibilism.

Part 3 is devoted to the Arminian accusations that Calvinism calls into question the goodness and justice of God. Chapter 6 is a discussion of the justice of Hell. Issues of predestination to damnation were revealed in chapter 3 to be of central concern in the debate. To that end, chapter 6 responds to the suspicion that Hell is unjust because it is disproportional: infinite punishment for finite sins. I rebut the charge by way of three suggestions: that the damned may continue to sin in Hell and thereby accrue more punishment; that human beings can be plausibly considered guilty for counterfactual sins, sins they would have done, and these are infinite; and I also defend the Anselmian suggestion that sins against an infinite being are of infinite gravity.

Chapter 7 presents a Calvinist theodicy of Hell. If God determines everything, why would God determine a large portion of humanity to reject Christ and go to Hell? In this chapter I rely heavily on the work of Jonathan Edwards, and I suggest that the reason why God predestines many to such a fate may be on account of the a greater sense of God’s justice, power, and greater thankfulness and appreciation for their salvation the elect in Heaven have by being part of the elect few, and defend the view against objections.

Chapter 8 deals with the question of divine paternity. It is objected by certain of Calvinism’s critics that for God to deterministically predestine an individual to damnation is inconsistent with God’s status as father. In this chapter I motivate Calvin's suggestion that God is the father only of the elect. I argue that if God is the father of all, then he is often in a sort of emotional paralysis, for, in any conflict between human beings, paternal duty and affection precludes him from identifying with (supporting) the interests of one party over the other.

The conclusion, chapter 9, offers some suggestions for future research.

## Acknowledgements

Daniel Hill and Stephen McLeod supervised this thesis, and proved very able and helpful supervisors. Daniel was my primary, and I very much appreciate the detailed and dedicated attention he gave to drafts of the various chapters of this thesis—attention that often extended to lengthy correspondence.

Portions of this thesis have already seen print. Chapter 4 is an altered version of my 2017 paper, ‘A Modest Classical Compatibilism’, published in *Disputatio* (45): 265–285, and I retain copyright. Chapter 7 is a modified version of ‘Calvinism and the Problem of Hell’—my chapter in the book *Calvinism and the Problem of Evil*, edited by David E. Alexander & Daniel M. Johnson, and published in 2016 by Pickwick Publications, an imprint of Wipf and Stock. It is used here by permission of Wipf and Stock Publishers ([www.wipfandstock.com](http://www.wipfandstock.com)).

I am very grateful to the Arts and Humanities Research Council for a studentship which made these studies possible. I was awarded an AHRC Northwest Consortium Doctoral Studentship for the 2015–18 period (award ref. 1635247).

## Table of Contents

Abstract.....	2
Acknowledgements.....	4
PART 1—THE GROUNDWORK.....	8
Chapter 1—What is Theological Determinism?.....	8
1.1 Species of Determinism.....	8
1.2 Is Determination Entailment?.....	10
1.3 Is Determination Causation?.....	12
1.4 Stating Theological Determinism.....	15
1.5 Other Issues.....	18
Chapter 2—The Free-Will Debate.....	20
2.1 Compatibilism and Incompatibilism.....	20
2.2 Point of Clarification (i): The Sort of Possibility at Issue.....	20
2.3 Point of Clarification (ii): Derivative and Non-derivative Moral Responsibility.....	21
2.4 Point of Clarification (iii): The Nature of Decision.....	21
2.5 Point of Clarification (iv): Agent Causation.....	22
2.6 Point of Clarification (v): The Nature of Moral Responsibility.....	23
2.7 Libertarianism and Hard Determinism.....	26
2.8 Revision to the Terminology.....	27
Chapter 3—The Free-Will Debate and Reformed Theology.....	30
3.1 Introduction.....	30
3.2 The Luther-Erasmus Exchange.....	30
3.3 Soteriological Calvinism and Deterministic Calvinism.....	35
3.4 Luther on the Ability Requirement for Moral Responsibility.....	36
3.5 Calvin’s Doctrine of Predestination.....	37
3.6 Calvin’s Deterministic Convictions.....	40
3.7 Calvin’s Dual-Actor Principle.....	42
3.8 Calvin on God as Author of Sin.....	44
3.9 Calvin on the Ability Requirement for Moral Responsibility.....	45
3.10 Arminius and the Belgic Confession.....	47
3.11 The Remonstrants and the Synod of Dort (1618–1619).....	50
3.12 Molinism.....	53
3.13 The 1646 Westminster Confession.....	55
3.14 The Determinism of the Westminster Confession: (i) God’s Decree.....	56
3.15 The Determinism of the Westminster Confession: (ii) ‘Coming Most Freely’.....	58
3.16 The Determinism of the Westminster Confession: (iii) Divine Aseity.....	59
3.17 The Westminster Confession Opposed to Determinism?.....	60
3.18 Wesley, Channing, and Westminster.....	62
3.19 John Wesley.....	63
3.20 William Channing.....	65
3:21 Summing Up.....	67
PART 2—THEOLOGICAL DETERMINISM AND FREE WILL.....	68
Chapter 4—The Ability Argument.....	68
4.1 Introduction.....	68
4.2 Analysis of modals.....	71
4.3 Analysis of ‘could have done otherwise’.....	73
4.4 Applying the Kratzer semantics to ‘could have done otherwise’.....	76

4.5 Frankfurt and other considerations.....	82
4.6 Conclusion.....	84
Chapter 5—The Manipulation Argument.....	85
5.1 Introducing the Argument.....	85
5.2 Mele’s Zygote Argument.....	87
5.3 Pereboom’s Four-Case Argument.....	89
5.4 The Hard-Line Soft-Line Distinction.....	93
5.5 Is Manipulation is the Problem?.....	97
5.6 Bignon’s Response.....	98
5.7 The Uncertainty Response.....	100
5.8 A Manipulation Argument for Compatibilism.....	102
PART 3—THEOLOGICAL DETERMINISM AND GOD’S GOODNESS.....	106
Chapter 6—The Justice of Hell.....	106
6.1 Introduction.....	106
6.2 What is Hell?.....	107
6.3 Lewis Carroll.....	109
6.4 The Proportionality Principle.....	110
6.4 Can finite sins merit only finite punishment?.....	113
6.5 The Central Argument.....	115
6.6 The Continuing-Sin Hypothesis.....	116
6.7 Denying FF.....	119
6.7.1 <i>More detail</i> .....	119
6.7.2 <i>Pan the Hedonistic Immortal</i> .....	119
6.7.3 <i>Partial and Fully Infinite Degrees of Punishment</i> .....	121
6.7.4 <i>Some Objections</i> .....	121
6.7.5 <i>The Significance of the Falsity of FF</i> .....	123
6.8 Against Our Sins Being Finite in Severity.....	125
6.8.1 <i>Anselm on the Status Principle</i> .....	126
6.8.2 <i>Jonathan Edwards on the Status Principle</i> .....	126
6.8.3 <i>Analysing the Status Principle</i> .....	127
6.8.4 <i>The Model in Greater Detail</i> .....	130
6.8.5 <i>The Case of Phanuel</i> .....	132
6.8.6 <i>Attacks on the Status Principle: False</i> .....	135
6.8.7 <i>Attacks on the Status Principle: Incoherent</i> .....	136
6.8.8 <i>Attacks on the Status Principle: Useless</i> .....	138
6.8.9 <i>Summary</i> .....	141
6.9 Against Our Sins Being Finite in Number.....	141
6.9.1 <i>Blameworthiness for what you would have done</i> .....	141
6.9.2 <i>Counterfactual Wrongs and Characters</i> .....	142
6.9.3 <i>Counterfactual Wrongs on Molinism</i> .....	143
6.9.4 <i>Counterfactual Wrongs on Theological Determinism</i> .....	144
6.9.5 <i>Objections Regarding the Use of Infinite Counterfactual Wrongs</i> .....	145
6.10 Conclusion.....	147
Chapter 7—Calvinism and the Problem of Hell.....	148
7.1 Introduction.....	148
7.2 The Problem of Hell.....	150
7.3 Calvinist Theodicy.....	152
7.4 Calvinist Theodicy Improved.....	155
7.5 Worthwhileness and Partiality.....	160
7.6 Two Objections.....	162

7.7 Conclusion.....	164
Chapter 8—Divine Love and Paternity.....	165
8.1 Introduction.....	165
8.2 The Talbott-Jordan Exchange.....	165
8.3 Flatness in Love.....	166
8.4 Deep Attachment.....	169
8.5 Does Deep Attachment Require Inequality in Love?.....	172
8.6 Romance and the Church.....	173
8.7 Argument from Incompatible Interests.....	174
8.8 Argument from Expressively Successful Deep Attachments.....	177
8.9 Summarising the Case Against UDP.....	179
8.10 The Case Against PDP: Creation and Paternity.....	180
8.11 The Case Against PDP: Scripture.....	182
8.12 God’s Love for the Reprobate.....	183
Chapter 9—Conclusion.....	185
Bibliography.....	187

# PART 1—THE GROUNDWORK

## Chapter 1—What is Theological Determinism?

### 1.1 Species of Determinism

This thesis is a defence of theological determinism. So what is theological determinism? Reflection on the term indicates that it is a species of determinism. Very well. What, then, is determinism?

Charlotte Werndl defines it as follows:

A system is deterministic just in case the state of the system at one time fixes the state of the system at all future times. (2017: 669)

But this is not a good definition of determinism. For one thing, the condition suggested is not necessary. Determinism can obtain without it. Consider a world,  $w_1$ , where there were no laws of nature, so no past state of affairs entails (relative to laws of nature) any future state of affairs, or even makes it more or less likely. But God (necessarily) exists in  $w_1$ , and God (necessarily) causes, directly, every state of affairs in  $w_1$ 's history moment by moment. This would be a world which is clearly deterministic, but not one that would satisfy the above definition. It is a world where the future is determined, but not determined by the past: it is a world where everything (apart from God) is determined by God.

This might seem an unfair objection. One might say that the above author wasn't offering a definition of determinism *per se*, but really a definition of a certain species of determinism. Nomological determinism, perhaps. Well, if that is so, then it is better if that is made clear. But does the definition given work as a definition of nomological determinism? 'Nomological' here should be understood to mean 'relating to the laws of nature'. The thought behind nomological determinism is that the *laws of nature* will determine a unique future. Suppose we assembled a complete description of the natural world at the present time,  $t_1$ . If nomological determinism is true, then, relative to the laws of nature, only one future can possibly succeed  $t_1$ . This seems to capture well what Werndl was getting at, and we might therefore offer to define nomological determinism (ND) as follows:

(ND<sub>1</sub>) A world,  $w_x$ , is nomologically deterministic iff there is a time  $t_x$  in  $w_x$  such that a complete description of  $t_x$  in conjunction with the laws of nature at  $w_x$  entails the state of  $w_x$  at any time  $t_y$  later than  $t_x$ .

But this definition is not quite adequate. It can be counterexampled in the following way, and thereby shown to be insufficient for ND. Suppose there is a universe with a history of 100 moments. The first moment,  $t_0$ , given the laws of nature, is sufficient to fix all moments after that *up to*  $t_{50}$ . What things will look like in  $t_{51}$  is not determined by  $t_{50}$ . There are two possible futures:  $t_{51a}$ — $t_{100a}$  and  $t_{51b}$ — $t_{100b}$ . There will therefore be two possible worlds: one for each future. Let  $w_a$  be the world which contains  $t_0$  to  $t_{100a}$  be  $w_a$ , and the world containing  $t_0$  to  $t_{100b}$ ,  $w_b$ . Neither  $w_a$  or  $w_b$  are deterministic worlds. But they satisfy (ND<sub>1</sub>).  $t_{51a}$  and  $t_{51b}$  are both times, in  $w_a$  and  $w_b$  respectively, that are such that a complete description of that time in conjunction with that world's laws (and  $w_a$  and  $w_b$  have the same laws) entails the state of the universe at any later time.

Definitions of determinism offered by other authors are adequate to accommodate this possibility, however. Here is van Inwagen on determinism:

*Determinism* may now be defined: it is the thesis that there is at any instant exactly one physically possible future. (1983: 3)

Here is Kadri Vihvelin:

More precisely, determinism is the thesis that for every instant of time  $t$ , there is a proposition that expresses the state of the world at that instant, and if  $P$  and  $Q$  are any propositions that express the state of the world at some instants, then the conjunction of  $P$  together with the laws of nature entails  $Q$ . (2013: 3)

We therefore need to modify the definition:

(ND<sub>2</sub>) A world,  $w_x$ , is nomologically deterministic iff for any time  $t_x$  in  $w_x$ ,  $t_x$  is such that a complete description of  $t_x$  in conjunction with the laws of nature at  $w_x$  entails the state of  $w_x$  at any time  $t_y$  later than  $t_x$ .

This captures van Inwagen's suggestion. Vihvelin's suggestion is different. Her definition is time-invariant, as follows:

(ND<sub>3</sub>) A world,  $w_x$ , is nomologically deterministic iff for any time  $t_x$  in  $w_x$ ,  $t_x$  is such that a complete description of  $t_x$  in conjunction with the laws of nature at  $w_x$  entails the state of  $w_x$  at any other time  $t_y$ .

Vihvelin's account has the commitment that the laws of nature permit us to infer the state of the past from the state of the future. She frames matters this way because she thinks that 'it is a feature of the physical theories known to be deterministic' (2013: 239). I don't think it is wise for the nomological determinist to commit quite yet to the ability of the future to determine the past, so I will prefer (ND<sub>2</sub>) as the better formulation of ND. (ND<sub>2</sub>) looks weighty enough to generate all the problems that determinism is typically supposed to generate. At any rate, by insisting that *any* past time (with the laws of nature) must entail all future times, (ND<sub>2</sub>) avoids the problem with (ND<sub>1</sub>).

But how does it help us define theological determinism (TD)? It does offer us an idea of how relevant kinds of determinism are to be understood. Nomological determinism, if (ND<sub>2</sub>) is right, concerns the entailment of one set of facts concerned with what we do (facts about future times) by another set do with the laws of nature (the past in conjunction with the laws of nature). This suggests a general schema: X-ological determinism will be the entailment of one set of facts, typically concerned with human action, by another set of facts concerned with X.<sup>1</sup> Theological determinism, therefore, will be the entailment of a set of facts concerned with the actions we perform by a set of facts about God. This understanding of the determining in determinism therefore reduces it to the entailment relation. This conception is endorsed by Carl Hoefer: 'Logical entailment, in a sense broad enough to encompass mathematical consequence, is the modality behind the determination in "determinism."' (2016).

## 1.2 Is Determination Entailment?

I think this is fairly significant mistake in the understanding of what determination amounts to. I think that to determine something is to do more than to entail it, and determination is more than entailment. I think determining something demands some form of explanatory priority from the determiner, and entailment does not supply that. Here follow some considerations to help us see that the central relation at play in determinism must be more than mere entailment.

---

<sup>1</sup> Throughout, I use 'fact' in a loose sense that isn't at pains to distinguish between true propositions and the parcels of *terra firma* reality to which true propositions usually refer. Context and charity should help the reader decide in each case the most appropriate understanding.

(i) If determination is entailment, then everything determines itself. For entailment is a reflexive relation. P entails itself: P entails P. We are all, therefore, self-determiners. Every act we perform is self-determining. But no one, I think, wants to say that. This suggests determination is something different.

(ii) Entailment is non-symmetric (so sometimes A entails B and B entails A), but determining, as it functions in metaphysical discussion, is intuitively asymmetric. That is to say, intuitively, if A determines B, then B does not determine A. If the course the boat takes is determined by the way the captain steers the wheel, then it isn't the case that the course the boat takes determines the way the captain steers the wheel. To get an asymmetric relation, we will have to move beyond entailment.

There are also considerations from particular cases.

(iii) Some people think there are truths about the future now. In fact, they might think that there are comprehensive truths about the future. They might think that everything I will do on May the 1<sup>st</sup> 2030 is presently true. In other words, propositions of the form, <I will mow the lawn on May the 1<sup>st</sup>, 2030> are presently true. Perhaps they are persuaded by an argument from bivalence. Either it is true that you mow the lawn on May the 1<sup>st</sup>, 2030, or it is false. It must be one or the other. But many of those who think that there are present truths about future actions will deny being determinists. But this is curious if determining is entailment. If the threat of determinism is the threat of entailment, then entailment of their actions seems to hold. The present truth of <I will mow the lawn on May the 1<sup>st</sup>, 2030> entails that I will mow the lawn on May the 1<sup>st</sup>, 2030. They wouldn't, I believe, happily say they are alethic determinists about the future. They would deny being determinists of any sort. That suggests determining is more than entailing.

(iv) In a similar vein, lots of theistic philosophers think determinism is false. They think so because they believe it would rule out free will and moral responsibility. But they also think that God has exhaustive foreknowledge of the future. God therefore knows every action every human being will perform in the future. And if God knows that *p*, then it follows that *p*. So, suppose that God foreknows that <Jones will mow the lawn on May the 1<sup>st</sup>, 2030>. The fact that God knows it entails that Jones will do it. But, again, even though such philosophers grant that God's knowing entails that the relevant action will happen, they will deny being determinists—of any sort. They won't even grant that they are 'foreknowledge determinists'. They will deny that God's foreknowledge determines. This, again, suggests that determination must be something more substantial than mere entailment.

### 1.3 Is Determination Causation?

Perhaps we might say that determination is causation? The causation relation is intuitively irreflexive and asymmetric. It also handles the alethic and foreknowledge cases. Truths about the future (whether freestanding or known by God) do not cause the free actions that those truths describe. So, perhaps it is the spectre of causation that those concerned to deny determinism whilst accepting entailment are eager to banish. This is plausibly Jonathan Edwards' understanding:

By determining the Will, if the phrase be used with any meaning, must be intended, causing that the act of the Will or choice should be thus, and not otherwise: and the Will is said to be determined, when, in consequence of some action, or influence, its choice is directed to, and fixed upon a particular object. As when we speak of the determination of motion, we mean causing the motion of the body to be in such a direction, rather than another. (1977 [1754]: 141)

To determine something, for Edwards, is to cause it to be some way rather than another. But this is not an ideal account of what it is to determine. The chief problem with it is that it rules out the possibility of non-causal determination. But a few thinkers have either felt inclined to be determinists of a non-causal type, or have at least advocated the possibility of non-causal determination.

Stewart Goetz, for instance, has defended the notion of teleological determination. Insisting, along with Anscombe (1957) and Wittgenstein,<sup>2</sup> that reasons are not causes, and therefore that reasons-explanation is a different kettle of fish from causal explanation, he moves that we can therefore envisage two ways in which an action can be determined: determined by reasons, and determined by causes. Although he believes that human actions are determined in neither manner, he claims he finds it more plausible to believe that human actions are teleologically determined than causally determined. Goetz writes, 'were I to become convinced that noncausal libertarian agency is either impossible or too problematic, I would embrace some form of noncausal teleological compatibilism in which our free actions, though determined, are ultimately explained teleologically by purposes.' (2008: 3–4).

Hugh McCann holds to a teleological determinism of precisely that sort, though he holds that the reasons or purposes that determine human actions are God's, not ours. He agrees with Goetz that 'we can give noncausal explanations of decisions in terms of the agents' reasons without fear of invoking an underlying causal explanation.' (1995: 584). However, McCann wants to preserve the

---

2 See Queloz 2017 for historical overview and commentary on Wittgenstein's position.

principle of sufficient reason (PSR). He therefore wants it to be the case that there is sufficient explanation for why it is that an agent freely acts one way rather than another. He thinks that such an explanation is to be found in God's mind:

There is, however, a solution which if successful would allow us to have things both ways [preserve free will and the PSR]. Suppose that God, as creator, is directly responsible for each of my decisions. If so, then even though my decision to vacation in Colorado was not determined by the rest of my nature, it still has an accounting—an accounting in terms of God's plans, of the good He sees in my deciding as I do. That is to say, what fully accounts for my decision is not my reasons for it, but God's. (1995: 585)

He appears to grant that 'whether we choose to call it nomic [nomological] or not, we still seem to have a brand of determinism' (1995: 593), viz. divine determinism, but insists that 'God's creative determination of my decisions does not rule out their being free' (1995: 593), because God's determining our decisions is a fundamentally different kind than the causal determination that was held to be threat to free will—as I read McCann, God teleologically determines our free choices. McCann understand God's determining in the following way:

[God's] relationship to us is not analogous to that of the puppeteer to his puppet—which would indeed destroy our freedom—but rather to that of the author of a novel to her characters. The characters do not exist as an event-causal consequence of anything the author does. Rather, their first existence is in her creative imagination, and they are born and sustained in and through the very thoughts in which she conceives them, and of which they are the content. (2005: 146)

God's determination is outside the causal chain. It is a determination of reasons, purposes, or teleology.

I'm not sure I believe that reasons-explanation is not causal explanation, but we can put that point to the side. It would be unfortunate if the definition of 'determinism' committed one to the falsity of views like Goetz's and McCann's.

It should also be noted that there are relations which, though not causal, might nevertheless be considered to be fit candidates to play a determining role. Much talk has been made in recent years of the 'grounding' relation. This is supposed to be a form of ontological dependence that can play a large and varied role. Much of our 'in virtue of'-talk is held by some to refer to this grounding relation. 'The wall has the shape it does by virtue of the placement of the bricks' can be informatively rendered as 'The shape of the wall is *grounded* in the arrangement of the bricks'; 'By

virtue of being persons, we have a right to life' can be informatively rendered as 'Our right to life is *grounded* in our personhood'; 'the existence of the singleton set {Socrates} exists by virtue of the existence of Socrates' can be analysed as 'the existence of {Socrates} is *grounded* in the existence of Socrates'; and so on.<sup>3</sup> In all those cases we don't want to say the relation is causal—Socrates doesn't *cause* the existence of {Socrates}, nor does our personhood *cause* our right to life—it is a relation, it appears, of non-causal determination. (The truth-making relation, that purportedly holds between reality and a true proposition, also looks to be non-causally determining. The truth of <the Pope is male> appears determined by the Pope's being male.)

At any rate, whether it be grounding or teleological determining, these are putative cases of non-causal determination that a definition of 'determinism' should accommodate. If someone were to insist that all human action had God as a sufficient ground, or that every aspect of the human decision-making process is teleologically determined by some superbeing, but deny that there were any causally sufficient conditions in play, then they would look like they were a determinist. Perhaps they are muddle-headed, but it is better not to assume so. The determining relation must be taken to be broader than mere causation, therefore. Even if it is not.

So, we need a relation broad enough to encompass causing, teleologically determining, grounding and any other intuitively determining candidates. I think the best thing we can say is that determination should be taken to be a metaphysical 'in virtue of' relation, a form of metaphysical explanation. We use the 'in virtue of' locution to cover instances of causing ('the avalanche occurred in virtue of last night's snowfall'), teleologically determining ('he choose to  $\phi$  in virtue of the strength of his reasons to  $\phi$ '), and grounding ('the statue's shape obtains by virtue of the arrangement of the lump of clay'). We seem to see in all those cases a shared aspect or common nature to the relation that makes it intuitively correct to speak of them all as 'in virtue of' relations.

Now, I'm not sure that there are instances where we would use the 'in virtue of' expression to denote a relation of *non-metaphysical* explanation, but if there are then I don't want to be committed to thinking of those cases as determining cases. The way I understand the 'in virtue of' relation is such that the following inference is valid: <A holds in virtue of B>; therefore, <B is metaphysically prior to A>. This won't hold, I take it, if there are true non-metaphysical instances of 'in virtue of'.

It is important to note that the metaphysical 'in virtue of' cannot be identified with relations of metaphysical priority or posteriority. To return to the mereological case, it is intuitive to think that the parts are metaphysically prior to the whole, such that for any P that is a part of a whole W, P is metaphysically prior to W. Let's assume our intuitions are right here. But it doesn't follow that

---

3 See Raven 2015 for an overview.

every aspect of the whole holds *in virtue of* every part. Suppose the whole in question is a statue of a soldier wielding a curved sword. All the parts of that statue are prior to the whole, but the parts of that statue that are located in the feet are not those parts in virtue of which the sword is curved. It is the parts of the statue located in the sword by virtue of which the sword is curved.

Of course, the metaphysical ‘in virtue of’ is an irreflexive and asymmetric relation. Nothing is the case in virtue of itself, nor can A hold in virtue of B whilst B holds in virtue of A. This ensures that the metaphysical ‘in virtue of’ avoids the problems afflicting the standard understanding of determining as entailment. It also provides a satisfactory explanation for why it is that (libertarian) foreknowledge theists and those who believe that there are determinate future truths about our actions do not consider themselves determinists: it is because in neither case is God’s foreknowledge or truths about the future that in virtue of which the future action will occur. They might entail it, but they aren’t what will bring about your future action.

The metaphysical ‘in virtue of’ therefore appears the best candidate for identifying the determining relation in determinism. When I use the term ‘determining’, it should be understood hereon to refer to this metaphysical ‘in virtue of’ relation. So understood, I take the determining relation to be irreflexive, asymmetric, and transitive. One should also note that the determining relation should be taken to entail entailment. In other words, that it is not possible for A to determine B without it being the case that A entails B. Although the determining relation can’t be identified with entailment, it does imply entailment.

We are familiar, therefore, with what it is for one fact to determine another fact. That, I hope, is clear from the foregoing illustrations and explication. Thus armed, we can now amend ND<sub>2</sub> in order to get a better understanding of nomological determinism:

(ND<sub>4</sub>) A world,  $w_x$ , is nomologically deterministic iff for any time  $t_x$  in  $w_x$ ,  $t_x$  is such that the set of facts that includes (i) all facts in  $w_x$  that hold at  $t_x$ , and (ii) the laws of nature at  $w_x$  (strictly redundant given (i)), determines the state of  $w_x$  at any time  $t_y$  later than  $t_x$ .

This evidently raises concerns about free will and moral responsibility, for any decision we perform will be at a time, and, if ND<sub>4</sub> is true of our world, then our decisions, and every aspect of our decision-making processes, will be fully determined by prior facts in our universe’s history.

## 1.4 Stating Theological Determinism

All well and good. But how, then, are we to define *theological* determinism? Here is one way:

(TD<sub>1</sub>) The set of all divine facts determines the set of all non-divine facts.

Divine facts are simply facts that make reference to, or somehow involve, God. The set of all divine facts would include God exemplifying his properties and God's actions. It would include God's foreknowledge. Even something like 'God's being pleased that Jones decided to  $\phi$  at  $t_x$ ' can be included. Although it entails that Jones decides to  $\phi$ , it doesn't determine that fact. However, there will be some other fact (or set of facts) about God that do, if the theological determinist is right.

There are some problems with TD<sub>1</sub>, however. The big one concerns abstract objects. The relation of God to abstract objects is a matter of some dispute among theists.<sup>4</sup> Abstract objects are often held to be necessarily existent, immaterial, non-spatial entities. In this category lie things like numbers, properties, moral values, and sets (though sets are, when they hold contingently existing members, contingently existent). Some philosophers will insist that God determines even the nature and existence of abstract objects (indeed, I fancy myself amongst that number); others (such as Craig (2016)), seeking to preserve divine sovereignty, will conclude that abstract objects aren't as real as we think they are; but some will say the realm of abstracta includes necessarily existing entities that God had no hand in creating or affecting. van Inwagen (2009) is such a one.

We should try to accommodate views such as van Inwagen's. But TD<sub>1</sub> does not do that. Necessarily existing abstract objects are non-divine, but God does and cannot determine either their existence or their nature, on van Inwagen's view. It is a bad consequence that anyone taking van Inwagen's view is *ipso facto* precluded from being a theological determinist. We can get around this by amending TD<sub>1</sub>:

(TD<sub>2</sub>) The set of all divine facts determines the set of all non-divine concrete facts.

Concrete entities stand opposed to abstract entities. Concrete objects are particular, causally efficacious, and rooted in *terra firma*. God himself is therefore a concrete object. But we don't want everything about God to be determined (that would imply circularity), hence TD<sub>2</sub>'s restriction to 'non-divine' concrete facts.<sup>5</sup>

---

4 See the edited volume, Gould 2014.

5 TD<sub>2</sub> still leaves open the question of God's relation to contingently existing abstract objects, such as sets of contingent things. However, I think that, for those who follow van Inwagen's view, God can only partially determine the existence of such objects: God will determine the existence of the contingent concrete objects which are members of these sets, but the necessary metaphysical law that generates the existence of these sets from the existence of the member or members is something that God could not have brought about, nor have any control

Some distinctions between types of theological determinism remain to be drawn, however. Consider what we might call ‘absolute determinism’:

(AD) For every fact, there is a fact that determines that fact.

Given how we defined ‘determining’, AD would issue in an infinite regress. That is surely problematic, and a good reason to reject AD. There must be some fact or facts that are undetermined, therefore. The naturalist, with his nomological determinism, would take the undetermined facts to be the existence and state of the universe at the first moment in time (and the laws of nature, if such laws are not determined by the initial state of the universe). The theist, on the other hand, with his theological determinism, will naturally take the undetermined fact or facts to be something concerned with God—at the very least, God’s existence.

There will therefore be different varieties of theological determinism that will take different aspects of God to be undetermined. How determined does the theological determinist want the set of divine facts to be? A very strong variety of TD would run as follows:

(TD<sub>3</sub>) God’s existence determines every concrete fact.

This view claims that the mere existence of the divine nature is sufficient to determine every other (concrete) fact, *including* facts about what God decides to do. Those who take this view would likely say there is only one undetermined fact: God’s being there. Everything else that happens is just the necessary, inevitable outworking of the divine nature. The striking commitment of this view is this: that God’s free decisions are also determined—determined by his nature. Some theists might take issue with that commitment. Some theological determinists might think that in order for God’s action’s to be free in the way that divine perfection requires, it must be that they are undetermined. I think such theological determinists should weaken TD<sub>3</sub> to

(TD<sub>4</sub>) The set of God’s decisions determines every non-divine concrete fact.

This is sufficient to have the whole history of the world, with all the human acts contained therein, from the noblest to the most wretched, determined by God’s decree, all the while leaving it open whether or not God’s own free choices are determined.

---

over, on van Inwagen’s view.

I am more sympathetic to TD<sub>3</sub> than merely TD<sub>4</sub>, but my defence of theological determinism here shall be, for the most part, equally applicable to either variety. TD<sub>2</sub> is the umbrella view that covers both.

## 1.5 Other Issues

One other important relation I believe one should be attentive to is what I shall call ‘fixing’. Suppose you have a situation where some fact F<sub>1</sub> is metaphysically prior to another fact F<sub>2</sub>, and F<sub>1</sub>’s being the case entails that F<sub>2</sub> will be the case, but F<sub>1</sub> does not determine (even partially determine) F<sub>2</sub>. In such cases, we can say that F<sub>1</sub> *fixes* F<sub>2</sub>. This issue is arguably of importance in the freedom and foreknowledge debate.<sup>6</sup> God’s foreknowledge of a future free action of mine *entails* that I perform that action, but does not determine it. Moreover, it is metaphysically prior to my act, or at least it appears to be so (doesn’t the present come before the future, not merely temporally, but in order of explanation?). God’s foreknowledge therefore fixes my future choice. One pertinent question is whether rejecting that God determines my free decisions is enough to preserve free will when it remains the case that my free decisions are fixed. This is a question libertarian foreknowledge theists must answer.

We must also remember that theological determinism is a claim distinct from necessitarianism. We can define necessitarianism as follows:

(N) For any fact that is the case, that fact is the case necessarily.

TD<sub>3</sub> looks like it entails N. God’s existence, even if not determined, obtains necessarily, in the view of most theists. God, as the greatest possible being, cannot merely exist contingently, by mere good fortune as it were. But God’s existence determines everything else given TD<sub>3</sub>, and if God necessarily exists and that existence determines everything else (including truths about what that existence determines), then it looks like everything else will hold necessarily as well. TD<sub>4</sub>, on the other hand, is not committed to N. It leaves open the possibility of divine properties that do not hold of necessity. In particular, the chief attraction of this view is, as one might expect, that it permits the theological determinist to hold that God’s decisions do not hold of necessity. And, if God’s decision,

---

<sup>6</sup> Indeed, the term ‘fixing’ was one inspired by discussions of this very issue, which often involve reference to the ‘fixity of the past’. See Fischer and Todd 2015 for an overview.

say, to create the world, holds contingently, then the world will also have its existence contingently, not necessarily. This might be thought a claim desirable to secure.

## Chapter 2—The Free-Will Debate

### 2.1 Compatibilism and Incompatibilism

Thus armed with sensible understandings of determinism, both nomological and theological, I proceed to central issues of dispute in the debate over ‘free will and determinism’ or, to use the traditional terminology, ‘liberty and necessity’ (van Inwagen, 1999: 341).

These are the central terms that need defining: ‘compatibilism’, ‘incompatibilism’, ‘libertarianism’, and ‘hard determinism’, although I will suggest revision to the terminology.

Compatibilism and incompatibilism are modal theses. We can define them as follows:

Compatibilism =df it is possible that an agent makes a decision, that every aspect of that decision-making process is determined, and that the agent is non-derivatively morally responsible for that decision.

Incompatibilism =df it is not possible that an agent makes a decision, that every aspect of that decision-making process is determined, and that the agent is non-derivatively morally responsible for that decision.

Here follow some points of clarification:

### 2.2 Point of Clarification (i): The Sort of Possibility at Issue

The sort of possibility referred to is metaphysical or broadly logical possibility. This is the sort of possibility that is considered as lying in between mere logical possibility (what the laws of logic permit) on the one hand and nomological possibility (what the laws of nature permit) on the other.<sup>7</sup> A state of affairs that is logically possible but broadly logically impossible would be something like a ball’s being red at every point and green at every point. We can see that it is something that cannot happen, but we also see that the impossibility of its happening isn’t grounded in the laws of logic (<‘red all over’(a) & ‘green all over’(a)> is not a logical contradiction), and neither is it grounded in

---

<sup>7</sup> Cf. Plantinga: compare ‘*David’s having travelled faster than the speed of light* and *Paul’s having squared the circle*. The former of these last two items is causally or naturally impossible; the latter is impossible in that broadly logical sense.’ (1974: 44).

the laws of nature (they don't seem to cover matters this fundamental). Consider that even if the laws of nature were different, it still wouldn't be possible for a ball to be red at every point and green at every point. The source of the impossibility of such a thing is therefore best described as being found in the 'laws of metaphysics'. Hence, metaphysical possibility.

### **2.3 Point of Clarification (ii): Derivative and Non-derivative Moral Responsibility**

The distinction between non-derivative and derivative responsibility can be understood through illustration. Suppose I decide to get recklessly drunk. In such an inebriated state, I might make many unfortunate decisions. Suppose I decide to vandalise someone's property. It would do no good to plead before the magistrate that I cannot be blamed for defacing the poor fellow's property because such actions were performed 'under the influence'—while I didn't have full control of my faculties. It will be responded that I was quite myself when I decided to get drunk in the first place, and I can thereby be held responsible for any mischief I decide to engage in as a result of the drunkenness I knowingly decided to inaugurate. I am therefore *derivatively* morally responsible for vandalising the gentlemen's house, and *non-derivatively* responsible for deciding to get drunk. My responsibility for the act of vandalising is somehow parasitic upon—derived from—my responsibility for getting drunk initially. Clearly, it is non-derivative moral responsibility that the free-will debate is chiefly concerned with.

### **2.4 Point of Clarification (iii): The Nature of Decision**

An explication of what it is to decide is not demanded by the definition, but is helpful regardless. I take a decision to be the forming of an intention. This understanding is standard. McCann states contemporary orthodoxy on the matter: 'Unlike intending, which is a state, deciding is an event. It is the mental act by which, in cases of fully deliberate action, reasons and intention are linked. Decisions are acts of intention formation, and so terminate in states of intending.' (1998: 133). It is also of note that it is intention *formation* that is understood to comprise decision-making. For *forming* an intention is different from *acquiring* an intention. To form an intention requires some sort of active contribution from the agent. To acquire an intention is simply to have it obtain in one's mind. One naturally associates acquiring an intention with passivity, though such passivity is not

essential to the idea. After all, by forming an intention one thereby acquires one. The converse is false, however. In brief, forming an intention is actional; acquiring one is not. See Mele (2017: 9–8) for discussion. Mention of agential contributions brings me to my next point.

## 2.5 Point of Clarification (iv): Agent Causation

I phrased the above definitions using the expression ‘every aspect of the decision-making process’. This was done deliberately, because there is debate amongst incompatibilists about where exactly to the indeterminacy has to go to get morally responsible decision-making. If I had said that compatibilism was the view that it is possible for *an agent’s decisions* to be determined yet for that agent to be non-derivatively morally responsible for them, then most incompatibilists would have agreed. That is because most incompatibilists think that morally responsible decision-making requires ‘agent causation’. Agent causation is a species of substance causation, and substance causation is typically placed in opposition to event causation. In brief, substances are the independent bearers of properties, and events are the exemplifications of a property at a time. Agent causalists hold that it is an inadequate account of what it is for an agent to decide to do something to make it a matter entirely of events. Suppose an agent’s deciding were a matter of event causation. What would the event be that caused the instantiation of the intention in the agent? It would be some psychological property or state. A certain set of beliefs or desires, perhaps. Or maybe, fundamentally, some sort of neurological property, such as an oscillation in brain waves or what have you. Agent causalists don’t like this sort of picture; they like to complain of ‘the disappearance of the agent’ (Pereboom 2014: 32). If all that happens when an agent decides is that one psychological event causes another, then it seems like the agent has vanished. Surely it isn’t the case (in morally responsible decision-making) that psychological processes cause my decisions; surely it is *I*, it must be *me*, that causes the intentions to form. And I am not an event. I am a substance. A simple argument of the form

(1) Non-derivative morally responsible decision-making requires the agent to form the intention in question.

(2) Event causalists hold that everything causally relevant to the formation of intentions can be understood in terms of events.

(3) If everything causally relevant to the formation of intentions can be understood in terms of events, then the agent isn’t forming an intention.

(4) Therefore, event causalism is incompatible with non-derivative morally responsible decision-making.

can be held to capture the basic thought.<sup>8</sup> Agent causalists believe that the agent *qua* agent must be responsible for the decision, and this requires him to cause as a substance—occurrent instantiations of properties aren't enough.

I am broadly sympathetic, and am inclined to hold to agent causation both as a matter of fact and as necessary for morally responsible decision-making. But I am not concerned to settle the issue as part of this thesis. One can be a theological determinist and either accept or reject agent causation<sup>9</sup>—it is a dispute orthogonal to the one I am engaged in. The point to be noted here is just this: that agent causalists hold morally responsible decisions to be *caused*—caused by the agent. If the agent causalist is an incompatibilist, then he will say that indeterminacy is required. But he will typically say that the indeterminacy must consist of this: that the *agent's causing of the decision* is undetermined. The agent must act as an undetermined determiner, a first mover, and the indeterminacy is thereby placed at the beginning of the process of decision-making, not in the middle.<sup>10</sup> The compatibilist agent-causalist, however, will hold that the agent's causing of the decision can be determined and this is no threat, in itself, to moral responsibility for the decision made, and in that way my definitions of compatibilism and incompatibilism accommodate agent-causal compatibilists and agent-causal incompatibilists.

## 2.6 Point of Clarification (v): The Nature of Moral Responsibility

I also use the expression 'moral responsibility'. What is that? We speak of morally responsible agents, and speak of people being morally responsible for certain actions or events. Of these two, I think we should take the latter as the more fundamental notion. Once we have what it is for someone to be morally responsible for some piece of behaviour, we can define a morally responsible person as 'someone able, by and large, to engage in morally responsible behaviour'. So what is it to be morally responsible for a piece of behaviour? Fischer and Ravizza have a good description of the phenomenon in question:

---

8 For an argument of this nature, see Franklin (2016: 1120–1121).

9 See Sehon 2016 for a defence of agent causation that doesn't require indeterminism.

10 Not every incompatibilist is happy with indeterminacy exclusively at the start of the process. Goetz (2008) complains loudly that is one's *decisions* that must be uncaused; it is no good at all if they are determined, even determined by the agent himself.

When we accept that someone is a moral agent, this [...] entails a willingness to adopt certain attitudes toward that person and to behave toward him in certain kinds of ways. (1998: 1)

What sort of ways?

When we regard someone as a responsible agent, we react to the person with a unique set of feelings and attitudes – for example, gratitude, indignation, resentment, love, respect, and forgiveness. (1998: 5)

Strawson calls such attitudes the reactive attitudes. They seem to be attitudes that it is only ever appropriate to direct towards persons, and in cases of attitudes like gratitude and indignation, only ever appropriate to direct toward somebody on account of a particular piece of behaviour from them. But such behaviour must issue in a morally responsible fashion. Strawson gives the following illustration:

If someone treads on my hand accidentally, while trying to help me, the pain may be no less acute than if he treads on it in contemptuous disregard of my existence or with a malevolent wish to injure me. But I shall generally feel in the second case a kind and degree of resentment that I shall not feel in the first. (1993 [1962]: 49)

That is because someone who steps on your hand accidentally while trying to help you has his moral responsibility for the act considerably diminished. He did not know, nor have any expectation (we can suppose) that his action would hurt you. To be sure, there is a sense in which they remain morally responsible: we are accountable to someone for unintentionally causing them pain—we might owe them some remuneration, say—but the point remains that diminished moral responsibility ('It wasn't really his fault!') leads to diminished indignation, that is to say, a diminishing of the relevant reactive attitudes.

The sort of reactive attitudes that are central to moral responsibility are arguably praiseworthiness and blameworthiness. To praise someone for what they have done and to blame someone for what they have done are the two central activities that moral responsibility centres around. When I talk of moral responsibility, these are the reactive attitudes that should be brought to the reader's mind. One might, if one wishes, replace 'non-derivatively morally responsible for that decision' in the definitions given, with 'non-derivatively praiseworthy or blameworthy for that decision'.

There is also the phenomenon of being obliged, or obligated. The language of duty is an important part of our moral discourse. We speak of what one ‘ought’ to do. But I don’t intend to include obligation or being obligated under the rubric of moral responsibility. This is because one can arguably obtain without the other. Suppose I promise to meet a friend at noon. However, someone poisons my morning coffee and I am rendered comatose until 1 o’clock. I was not morally responsible for missing my appointment with my friend. I could not be blamed for it. But I was arguably still under an obligation to meet my friend at noon, even when I was in a coma at 11:59. It is strange to say that the obligation vanishes when I become comatose, though one might say it. The more natural thing to say, it seems to me, is that I failed in my duty, but that I was not blameworthy for this failure. The unfortunate catatonia removes my responsibility for my failure in duty, it doesn’t remove the duty. For that reason, I think that obligation, and the conditions necessary and sufficient for that phenomenon, should be treated differently from the conditions necessary and sufficient for moral responsibility.

There are typically two conditions which are suggested as necessary for moral responsibility for some piece of behaviour (Rudy-Hiller, 2018): (a) the epistemic condition, and (b) the control condition. Absence of either diminishes (if not removes) responsibility for the behaviour.

We can understand the need for an epistemic condition by considering a certain type of excuse that we use to respond to accusations of fault. If someone unintentionally treads on Strawson’s hand and Strawson becomes indignant at this, then it is natural for the accused party to reply ‘But I didn’t expect that to happen’. Likewise, if I unwittingly run over my neighbour’s cat, then I would be likely to deal with my neighbour’s distraught accusations by replying ‘But I didn’t see it’. In both cases, the impression of blameworthiness is sought to be undermined by a claim to a *lack of knowledge*. Since this strategy is acknowledged as acceptable in principle—no one responds by saying, ‘I don’t care that you didn’t know. You are still just as blameworthy’—, this acceptance pushes us towards positing an epistemic condition on moral responsibility. One must be, in some sense, aware of what one is doing to be responsible for what one does.

Next there is the control condition. It is over this condition that the free-will debate takes place. In order to be responsible for one’s actions—in order to do them freely—one must have control over one’s actions. Or so the thought goes. I think that the control condition is really better understood as a cluster of putative conditions. The language we use here is very varied. We speak of our actions being performed ‘freely’ or ‘not freely’; of being ‘able’ to do something, and ‘unable’ to do it; of our choices being ‘up to us’, and their not being ‘up to us’; of being ‘in control’ of our actions, and our actions being ‘out of our control’.

However, I think that many of these can come apart. I am not alone in thinking this. John Martin Fischer (2010), for instance, thinks that ‘able’-talk isn’t strictly relevant to moral responsibility. It is natural to think there is an ‘able’ condition on moral responsibility because we use inability to defeat blame. We say things like ‘But I wasn’t able to do anything about it!’ in order to justify ourselves, which practice apparently presupposes that ability is necessary for moral responsibility. But Fischer thinks that what we are *really* interested in when we use such expressions is whether or not we have the right *control* over the relevant behaviour. Fischer therefore dismisses ‘able’-talk in favour of ‘control’-talk and goes on to identify the free-will dispute as being over what sort of control is required for moral responsibility. ‘Is it a control that requires the falsity of determinism, or is it not?’ (And Fischer says it is not.) We will return to Fischer’s position in ch. 4.

It isn’t my intention to sort through all our variegated moral-responsibility-talk, neither do I wish to offer a comprehensive account of the nature of moral responsibility. I will only comment on the relevant species of talk if it is brought up as a plank in an argument that moral responsibility is incompatible with determinism, or if it is otherwise germane to do so. As it turns out, there is an important species of argument that circles around ability ascriptions and a supposed tension between such ascriptions and determinism. Again, that matter will be covered in ch. 4.

## **2.7 Libertarianism and Hard Determinism**

With an understanding of compatibilism and incompatibilism, we are now in a position to understand libertarianism and what is called ‘hard determinism’.

Libertarianism =df Incompatibilism is true, and human beings are (in typical cases) morally responsible for their decision-making.

Hard Determinism =df Incompatibilism is true, and every aspect of our decision-making processes is determined.

Unlike the incompatibilism-compatibilism distinction, this distinction is not a modal one. The former distinction was about what is compossible with what. But this latter distinction is about what is actually the case. Is determinism actually true or not? Are we actually morally responsible or not?

The ‘in typical cases’ qualification present in libertarianism indicates that, of course, the libertarian does not think that every piece of decision-making we make is made in a morally responsible fashion. The libertarian is not committed to thinking the decisions we make when

sleeping or under compulsion are morally responsible decisions. Nothing revisionary is intended; there are everyday situations in which we are inclined to hold people responsible for their decisions and situations where we are not; the libertarian takes us to be indeed typically responsible in situations of the former sort, and adds that such responsibility is incompatible with every aspect of the decision-making process being determined.

Hard determinism gets its ‘hardness’ from its denial that human beings are morally responsible creatures. Unlike libertarians, hard determinists are therefore committed to a revisionary understanding of when human beings are responsible for what we do. The common-sense understanding is that human beings are frequently responsible for their behaviour; the hard determinist understanding is that they never are! Hard determinism is often associated with Derk Pereboom (2001). The hard determinist typically comes to his counterintuitive position because he is persuaded on the one hand by the arguments for incompatibilism, but is not persuaded on the other hand that indeterminism is present in our decision-making. Or at least not persuaded that whatever indeterminism is present in the decision-making process is present at the right point. This latter possibility suggests an improved understanding of hard determinism:

Hard Determinism\* =df Incompatibilism is true, and those aspects of our decision-making processes by virtue of which we would be morally responsible are determined.

## 2.8 Revision to the Terminology

I mentioned that I would revise the terminology. This is because ‘hard determinism’ doesn’t quite match up to the debate. We can leave it as we have defined it, but I move it should be replaced with ‘hard incompatibilism’ when it comes to the initial statement of the various views. The two most important questions at issue in partitioning off the various positions one can take in the free-will debate are these: (I) Are we morally responsible for our decisions? (II) Does moral responsibility (MR) for decisions require indeterminism? A typical tabulation of the responses looks like this:

	MR requires indeterminism	MR does not require indeterminism
We are morally responsible	Libertarianism	Compatibilism
We are not morally responsible	Hard Determinism	???

But ‘Hard Determinism’ is not suited to occupy the box in which it is placed. For one thing, the conjunction of the claims <moral responsibility requires indeterminism> and <we are not morally responsible> are perfectly consistent with the features of our situation that rule out our moral responsibility having nothing to do with determinism. Perhaps we are not morally responsible because we never satisfy the epistemic condition.

For another thing, the term ‘Hard Determinism’ could apply equally well to the bottom-right box. The conjunction of <moral responsibility does not require indeterminism> and <we are not morally responsible> can obtain just as well there too. Indeed, those who hold to such a position are probably likely to be determinist-friendly, simply because they think moral responsibility is compatible with it. One suspects that, while granting determinism, or being inclined to, they will rule out moral responsibility on other grounds that have nothing to do with determinism.

In short, the problem in both cases is that the ‘determinism’ in ‘hard determinism’ isn’t really doing any work as far as that table goes. If we want an answer to fit in that square (and that would be a helpful thing to want), then some revision is necessary. I propose the following:

	MR requires indeterminism	MR does not require indeterminism
We are morally responsible	Libertarianism	[Soft] Compatibilism
We are not morally responsible	Hard Incompatibilism	Hard Compatibilism

Replacing ‘Hard Determinism’ with ‘Hard Incompatibilism’ solves the issue.<sup>11</sup> The occupant of that box is indisputably incompatibilist, because he holds that moral responsibility requires indeterminism, and he is indisputably hard, because he denies human beings are morally responsible. Nothing is said about the grounds for our non-responsibility. Likewise for ‘Hard Compatibilism’.

At any rate, such are the terms, and that is what I take each of them to mean. There remain two issues with the revised table of views. First, there is no commitment to the truth or falsity of determinism (save in the case of libertarianism). I grant that, but I believe that must be considered, strictly, a separate issue. Secondly, the term ‘Compatibilism’ is most accurately used when used as I have defined it at the initial stage in this chapter, as the simple compatibility of moral responsibility with determinism. But it is often used more loosely than that because of the requirement to have a view to contrast with libertarianism. ‘Compatibilism’ so employed is taken as the conjunction of

11 Pereboom has already deployed the expression ‘hard incompatibilism’, and he takes it to be the claim that human beings are not responsible regardless of whether or not determinism is true (incompatibility on both fronts) (2001: xix). But he acknowledges that his use is revisionary (2001: 127), and I so I don’t feel much pressure to follow him on this point, and I find his suggested meaning too misleading in any case. I think a better term for his view would be ‘double incompatibilism’, for he thinks moral responsibility is incompatible twice over: with determinism *and* with indeterminism.

compatibilism and the belief that human beings are morally responsible. Such a position is of course something more than mere compatibilism, however. The best term for it, on account of its affirmation of man's moral responsibility, must be 'Soft Compatibilism'. However, it is difficult to go against the grain, and my language will suffer some imprecision of usage. I will sometimes write 'compatibilism' and mean by that 'soft compatibilism'. That is why I placed the 'Soft' in square brackets. The reader will have to judge from the context which expression is meant.

## Chapter 3—The Free-Will Debate and Reformed Theology

### 3.1 Introduction

So much for the terms and the philosophical debate about free will. But I am not concerned merely to deal with these doctrines in the abstract and to study their various entailments or implications. I want the issue to be historically concretised. In particular, I want to tie the issue to a debate that has continued for centuries in Protestant Christianity (especially in evangelical circles): the debate between Calvinists and Arminians.

On the usual understanding, Calvinists are theological determinists (and soft compatibilists), and Arminians are libertarians. Here is John Anderson on the first point:

It should be conceded at the outset, and without any embarrassment, that Calvinism is indeed committed to divine determinism: the view that everything is ultimately determined by God. I will not argue this point—it can be amply documented from representative Calvinist sources—but will simply take it for granted as something on which the vast majority of Calvinists and their critics agree. (2016: 204)

Though whether, historically, Calvinism has indeed been deterministic has been recently challenged. The volume *Reformed Thought on Freedom*, edited by van Asselt, Bac, and te Velde (2010), challenges the idea, as does Fesko (2014). I side very much with Anderson on this point, and in this chapter I shall argue from a historical survey of representative sources of Reformed thought that the Reformed tradition is a deterministic tradition. I shall begin at the dawn of the Reformation, with Martin Luther's *Bondage of the Will*, his response to Desiderius Erasmus's *Diatribes*, and I try to cover all salient developments up to the *Westminster Confession*, finally concluding with John Wesley's and William Channing's fierce attacks against Calvinism.

### 3.2 The Luther-Erasmus Exchange

Martin Luther (1483–1546) was born in Germany of humble stock, yet in 1512 was awarded a doctorate in theology, and received a chair in theology at the University of Wittenberg. His refusal

to bow to the authority of the papacy over the matter of indulgences led to his excommunication from the Catholic church in 1521, and thus began the Reformation. Once Luther had become established as one of the leaders of the Reformation, he became embroiled in various controversies. But the great controversy of relevance to my purposes is that which he had with Erasmus over the freedom of the will. Desiderius Erasmus (c. 1469–1536) was a famous Renaissance scholar, humanist, and wit. In 1516 he published his *Greek New Testament*, and thereby made no small contribution to the process of Reformation. But he refused to join with the Reformers. He took issue with much of what Luther said and how he said it. Among other things, he was not prepared to accede to Luther’s proclamations about the impotence of the human will.

The work that kicked off the dispute with Erasmus was Luther’s *Assertio* written (and published in its final, German version in March 1521) in response to a bull of admonition by the Pope, and in the thirty-sixth article of that *Assertio*, Luther penned the following:

Since the fall of Adam, or after actual sin, free will exists only in name, and when it does what it can, it commits mortal sin. This article ought to be clear enough from those that precede because St Paul says, in Romans 14:23, “Everything that is not of faith is sin.”

Though the earlier, Latin version of the *Assertio* from December 1520 contains stronger remarks:

I was wrong in saying that free choice before grace is a reality only in name. I should have said simply: “free choice is in reality a fiction, or a name without reality.” For no one has it in his own power to think a good or bad thought, but everything (as Wyclif’s article condemned at Constance rightly teaches) happens by absolute necessity.<sup>12</sup>

It was this earlier Latin version that Erasmus read and engaged with. He published his *Diatribes on Free Will* in 1524, intending to check Luther’s intemperate remarks about free will with something more reasonable. Erasmus defines free will as

[A] power of the human will by which a man can apply himself to the things which lead to eternal salvation, or turn away from them. (1969 [1524]: 47)

And his overall position is that though man’s will is, in man’s fallen state, a feeble thing, yet it has its own contribution to make to man’s salvation. He writes,

---

<sup>12</sup> As quoted by Marlow and Drewery in Erasmus (1969 [1524]: 13).

Nor in the meanwhile does our will achieve nothing, although it does not attain the things that it seeks without the help of grace. But since our own efforts are so puny, the whole is ascribed to God, just as a sailor who has brought his ship safely into port out of a heavy storm does not say: “I saved the ship” but “God saved it.” And yet his skill and labour were not entirely entirely useless. (1969 [1524]: 79)

And he objects to Luther’s view as follows. Speaking of Paul’s vigorous language in the epistles, he writes,

It seems to me difficult to associate the words “contest,” “crown,” “righteous judge,” “giving,” “fighting,” when all things happen from mere necessity with our will doing nothing, but merely passive. (1969 [1524]: 62)

And

What is the point of praising obedience if in doing good or evil works we are the kind of instrument for God that an ax is to a carpenter? But such a tool are we all if Wyclif is right. All things before and after grace, good equally with ill, yes even things indifferent, are done by sheer necessity. (1969 [1524]: 63–64)

It was therefore Luther’s appeal to a doctrine of absolute necessity that Erasmus zoned in on, and I suspect that Luther modified that part of *Assertio* because he felt it wasn’t quite to the point. Further down in the thirty-sixth article Luther says this:

Again, Moses says in Genesis 6:3 and Genesis 8:21, “Everything that the heart of man thinks and desires is only evil at all times.” Hearken to that, dear papists; Moses opens his mouth against you, what will you say in reply? If there is a good thought or will in men at any time, then we must accuse Moses of lying, for he calls all the times, all the thoughts, all the desires of the human heart evil. What kind of freedom is it that is inclined only to evil?

This makes it plain that Luther was grounding the supposed non-existence of free choice in the great wickedness of man’s heart. The doctrine that God determines all that comes to pass (though Luther appears to hold to it, judging by his endorsement of Wyclif) was not, at that point, what he wanted to draw attention to. Luther’s central contention was that, after man’s fall the corruption of man’s nature was so great that he is unable to will anything good. That is why Luther claims in the

1521 *Assertio* that when a fallen human being, such as you or I, do the best we can, morally speaking, all we will end up committing are mortal sins.

He argues for this position at length in his *Bondage of the Will*, published in 1525—his great response to Erasmus’s *Diatribes*. The work continues to garner for itself great praise. Godwell Chan says of it that ‘Luther’s work, a masterpiece, is irrefutable.’ (1996: 1) and Lee Gatiss opines that ‘If modern evangelicals have lost Luther’s clarity and faithfulness to Scripture on this issue of free will, we will have lost something very precious and foundational indeed.’ (2009: 203).

Luther writes in this acclaimed book that

‘free-will’ without God’s grace is not free at all, but is the permanent prisoner and bondsman of evil, since it cannot turn itself to good. (1957 [1525]: 104)

and

we do everything of necessity, and nothing by ‘free-will’; for the power of ‘free-will’ is nil, and it does no good, nor can do, without grace. (1957 [1525]: 105)

In such passages it looks like Luther reaffirms that our actions are necessitated, and that this is grounded in our evil, fallen nature. That’s why we don’t have ‘free will’. But he doesn’t merely affirm that the absence of free will follows from our wicked nature, however. He also at points reaffirms the Wycliffian doctrine that all things happen of necessity, and this on account of the nature of God. Luther writes against Erasmus,

For if you hesitate to believe, or are too proud to acknowledge, that God foreknows and wills all things, not contingently, but necessarily and immutably, how can you believe, trust and rely on His promises? (1957 [1525]: 83–84)

and

It is, then, fundamentally necessary and wholesome for Christians to know that God foreknows nothing contingently, but that he foresees, purposes, and does all things according to His own immutable, eternal and infallible will. This bombshell knocks ‘free-will’ flat, and utterly shatters it (1957 [1525]: 80)

This looks like a forthright statement of theological determinism: everything that occurs has been determined, or ‘necessitated’, by the intentions and purposes of God. Luther therefore appears to think there are two sources which are sufficient to demonstrate that there is no ‘free will’: (i) man’s native depravity and (ii) the infallible decrees of God.

But later remarks Luther makes are in tension with the determining power of the second of the those sources. In Luther’s *Table Talk*, we find him saying the following about ‘free will’:

I confess that mankind has a free will, but it is to milk kine, to build houses, etc., and no further; for so long as a man is at ease and in safety, and is in no want, so long he thinks he has a free will, which is able to do something; but when want and need appear, so that there is neither meat, drink, nor money, where is then free will? It is utterly lost, and cannot stand when it comes to the pinch. (1857: 120)

But this appears to undermine his earlier claim that the decrees and plans of God were sufficient to necessitate or determine everything. For God’s decrees extend no less to the milking of kine and the building of houses than they do to the spiritually pleasing to God (Luther insisted above that ‘all things’ happen according to God’s immutable will). How, then, can man be free even to milk kine? He will milk kine because God has decreed that he will milk kine. Therefore, according to the logic Luther presented before against Erasmus, man will not be free in his milking of kine.

Luther here appears to be drawing a distinction between spiritual matters and non-spiritual matters. Man has free will in the latter case (milking kine, etc.), it appears, but not in the former (trusting in Christ, etc.). That Luther has in mind such a distinction is borne out by other remarks in his *Table Talk*:

This is my absolute opinion: he that will maintain that man’s free will is able to do or work anything in spiritual cases, be they never so small, denies Christ. This I have always maintained in my writings, especially in those against Erasmus, one of the learnedest men in the whole world (1857: 119–120)

From such remarks one again gets the impression that Luther would be happy to grant man free will in non-spiritual cases—only to hold to free will in spiritual matters involves one in grave theological error.

### 3.3 Soteriological Calvinism and Deterministic Calvinism

The *Augsburg Confession* of 1530, drawn up by Philipp Melanchthon (1497–1560), and approved of by Luther, captures the convictions of the early Lutheran church, and it too stresses the importance of the distinction between spiritual and non-spiritual matters vis-à-vis free will. Here is content from Article 18, ‘Of Free Will’:

Concerning free will, they [the Lutherans] teach that man's will hath some liberty to work a civil righteousness, and to choose such things as reason can reach unto; but that it hath no power to work the righteousness of God, or a spiritual righteousness, without the Spirit of God; because that the natural man receiveth not the things of the Spirit of God (1 Cor. ii. 14). But this is wrought in the heart when men do receive the Spirit of God through the Word. (Schaff, 1977: 18)

We are therefore left with two distinct suggestions about what it is that cancels out human free will: it is either (i) man's wicked fallen nature, which determines him to refrain (at least) from choosing the spiritually good, or (ii) God's decrees, will, and foreknowledge, which determines every event in history, including every action of man. They both appear to be sources that determine or necessitate human action, and thereby rule out libertarian free will. But by the time of the Augsburg Confession, it appears that the emphasis was given to (i) over (ii).

This distinction between human action being necessitated by bondage to sin or by being the effect of divine decree has led to two sorts of Calvinism being discussed. Daniel Johnson has accordingly distinguished between what he calls ‘Calvinist soteriology’ and ‘Calvinist determinism’. Calvinist soteriology is the claim that ‘fallen man is unable to turn to God with saving faith, because man is *unwilling* to turn to God (and is therefore responsible for his rebellion). Every believer is infallibly brought to faith, sustained in faith, and sanctified by the omnipotent power of the Holy Spirit.’ (2016: 20), and Calvinist determinism being the claim that ‘God is in control of everything, and has from eternity ordained all that has come to pass and will come to pass. At the same time, human beings are genuine agents and are responsible for their actions.’ (2016: 21).

He is not the only one to draw this distinction. William Cunningham has written about what he terms the *philosophical* doctrine of necessity and compared it to what he calls the *theological* doctrine of necessity. The philosophical doctrine of necessity is effectively theological determinism, and the theological doctrine consists of the affirmation that

The necessity, or servitude, or bondage, which [the early Reformed] ascribed to the will of fallen man, consisted in the loss of the liberty [...], and in the actual prevailing tendency of his moral nature to evil because of the depravity which had overspread it, so that he could no longer will good but could only will evil. The liberty which they thus ascribed to man in his original condition, they regarded as entirely lost by the fall, and as having now no existence in men in their natural condition, or until restored, in some measure, by divine agency in regeneration. (1862: 505)

Cunningham therefore laments the

injurious tendency and consequences of this assumed identity or necessary connection of the two doctrines,—the theological and philosophical. It tends to throw into the background the true scriptural, theological doctrine of necessity,—the doctrine of the servitude or bondage of the will of fallen man,—man as he is,—to sin because of the depravity which has overspread his moral nature. (1862: 514)

There is indeed an important difference between these two doctrines. But I intend to show here that the Reformed tradition, as exemplified in its great confessions and the theology of its leading figures, is indeed committed to more than merely the theological doctrine of necessity; it is committed to the philosophical doctrine of necessity, to the view that God's decrees necessitate all things, that is to say, to theological determinism.

### **3.4 Luther on the Ability Requirement for Moral Responsibility**

But before we proceed with that, we would do well to note how Luther responded to Erasmus's earlier complaint, 'What is the point of praising obedience if in doing good or evil works ... All things before and after grace, good equally with ill, yes even things indifferent, are done by sheer necessity[?]' (63–64). It appears Erasmus believes that some sort of ability to do otherwise must be present for the sort of praise and blame we associate with moral responsibility to obtain. Luther's response to this sort of objection to his position is very striking:

*'Who' (you say) 'will try and reform his life?' I reply, Nobody! Nobody can! God has not time for your practitioners of self-reformation, for they are hypocrites. The elect, who fear God, will be reformed by the Holy Spirit; the rest will perish unreformed. [...]*

‘Who will believe’ (you say) ‘that God loves him?’ I reply, Nobody! Nobody can! But the elect shall believe it; and the rest shall perish without believing it, raging and blaspheming, as you describe them. (1957 [1525]: 99)

Luther wholeheartedly embraces the inability of man to do otherwise than he in fact does. But also insists, on account of plain fact that Christian doctrine states man’s guilt for his unbelief, that man is indeed responsible for not believing. This sort of objection to the Reformed position will occur again and again. We will see it in Calvin’s dispute with Pighius and Georgius, and we will see it from the pen of Wesley. For the most part, the later Reformed were not willing to fully endorse the extremity of Luther’s rhetoric in this matter.

### **3.5 Calvin’s Doctrine of Predestination**

The next great figure we encounter in Reformed history is John Calvin (1509–64). Calvin is famous for his doctrines of predestination; in particular, his account of election and predestination. He did not invent the terms, but the account he gave of them is widely considered, rightly or wrongly, the distinctive feature of his thought. Calvin summarises matters in this way:

By predestination we mean the eternal decree of God by which He determined with Himself whatever He wished to happen with regard to every man. All are not created on equal terms, but some are preordained to eternal life, others to eternal damnation. Accordingly, as each has been created for one or other of these ends, we say that he has been predestinated to life or to death. (*Institutes* 3.21.5)

We can see that Calvin’s convictions expressed here naturally comport well with the supposition that Calvin was of a theologically deterministic mindset. The opening sentence is not readily squared with a libertarian perspective. If God has the power to determine whatsoever he wishes to happen to any man, then God has the power to, one presumes, bring it about that a man should fall into sinful patterns of behaviour, or that a man should perform a great and noble act of self-sacrifice. But how can God have such control over the morally responsible actions of his creatures if libertarianism is true? He cannot cause or determine their decisions, because, on the incompatibilist assumptions of the libertarian, that would make them unfree—actions for which the agent is not morally accountable. But if compatibilism is true, then it appears quite possible for God to deterministically cause all of man’s actions yet for him to be accountable for them.

Calvin also mentions that some are predestined to eternal life, and others to eternal damnation. These are the decrees of *election* and *reprobation*, respectively. God elects an individual when he decides, and brings it about, that that individual will meet the conditions for salvation, and therefore enter into Heaven upon death. God reprobates an individual when he decides, and brings it about, that that individual will fail to meet the conditions for salvation, and therefore enter into Hell upon his death.

What is striking about Calvin's doctrine of election and reprobation is that he believes that both are *unconditional*. He denies, firstly, that prescience (that is, foreknowledge) is the ground of election or reprobation:

The predestination by which God adopts some to the hope of life and adjudges others to eternal death, no man who would be thought pious ventures simply to deny; but it is greatly caviled at, especially by those who make prescience its cause. We, indeed, ascribe both prescience and predestination to God; but we say that it is absurd to make the latter subordinate to the former. (*Institutes* 3.21.5)

Election is not therefore based on what God foreknows about an individual. God does not peer through the corridor of time, and, because he perceives that a person will be receptive to the gospel in the future (or some possible future), decide to predestine that person to receive it. But, more than that, election is not based on anything to do with the elected person at all. It based grounded simply in God's free pleasure. Here is more from Calvin on the reasons why God decides to bestow electing grace:

If you ask the reason[,] the apostle gives it, "For he saith to Moses, I will have mercy on whom I will have mercy, and I will have compassion on whom I will have compassion" (Rom 9:15). And what, pray, does this mean? It is just a clear declaration by the Lord that He finds nothing in men themselves to induce Him to show kindness, that it is owing entirely to His own mercy, and, accordingly, that their salvation is His own work. (*Institutes* 3.22.6)

Commenting on Ephesians 1:4–5, which reads as follows,

According as he hath chosen us in him before the foundation of the world, that we should be holy and without blame before him in love: Having predestinated us unto the adoption of children by Jesus Christ to himself, according to the good pleasure of his will (AV)

Calvin comments on Paul's last expression there, 'according to the good pleasure of his will', giving it the following interpretation:

Then, if a higher cause is asked, Paul answers that God so "predestined," and predestined "according to the good pleasure of his will." By these words he overturns all the grounds of election which men imagine to exist in themselves. For he shows that whatever favours God bestows in reference to the spiritual life, flow from this one fountain, because God chose whom He would and, before they were born, had the grace which He designed to bestow upon them set apart for their use. (*Institutes* 3.22.2)

In Calvin's mind, God's decision to elect individuals to salvation is therefore based on nothing in the individuals themselves. No condition or feature they possess moves God to elect them instead of his refraining from doing so. In that sense, election is unconditional.

One should bear in mind the distinction between *election* and *salvation*, however. Salvation is that conversion from death to life—from alienation from God to union with him—of which Scripture speaks. God's *electing* someone is his decision to bring it about that they are saved. Salvation can therefore be conditional even though election is unconditional: God simply unconditionally decides that an individual will meet the conditions for salvation. And, indeed, in Protestant theology, salvation is typically held to be conditional: faith in Christ being that condition.

Not merely is *election* unconditional in Calvin's thought, however, so is *reprobation*.

At last, [Paul] concludes that God hath "mercy on whom he will have mercy, and whom he will he hardeneth" (Rom 9:18). You see how he refers both to the mere pleasure of God. Therefore, if we cannot assign any reason for His bestowing mercy on His people but just that it so pleases Him, neither can we have any reason for His reprobating others but His will. When God is said to visit in mercy or harden whom He will, men are reminded that they are not to seek for any cause beyond His will. (*Institutes* 3.22.11)

Another quote to the same effect:

Those, therefore, whom God passes by, He reprobates, and that for no other cause but because He is pleased to exclude them from the inheritance which He predestines to His children. (*Institutes* 3.23.1)

Therefore, for Calvin, God's decision to bring it about that an individual is damned is likewise not grounded in anything in the reprobated individual. It is God's free pleasure.

Now, unconditional election and reprobation are what we would expect if Calvin's perspective was fundamentally a theologically determinist one. 'If God determines everything,' one might say, 'why would election be based on anything in man or that man does? The only things that are in man or that he does are the things that God has determined him to be or to do.' And it is surely right that conditional election blends far more naturally with a libertarian perspective. God cannot determine free actions on that theory, and so if salvation involves a free decision, God can only save those who would satisfy that condition. It would therefore makes perfect sense, on the libertarian view, for God to elect on the basis of foreknown free decisions to accept the offer of salvation. Calvin, in roundly eschewing that perspective, therefore finds himself in natural alignment with the theological determinists.

But it isn't as strong a proof as one would wish of Calvin's determinism. After all, returning to Daniel Johnson's distinction above between Calvinist soteriology and Calvinist determinism, it might be argued that Calvin's belief in unconditional election and reprobation doesn't bespeak determinist convictions, but merely a concern to remove entirely human merit from the equation—arguably one of the distinctive claims of 'Calvinist soteriology' (if man is so bound up in moral darkness that he cannot lift even a mental finger to will the good, then it plausibly follow that he incapable of doing anything whereby he might merit salvation). Man might therefore have free will as the libertarian understands it (in non-spiritual matters), but such free will would play no part in election or reprobation.

It is certainly true that Calvin wished to rule out the idea that human beings could in any way merit their salvation or election. One of the reasons he gave for insisting that foreknowledge could not ground election was to block of the 'common imagining' that God predestined according to foreseen merits. Calvin writes,

For they commonly imagine that God distinguishes between men according to the merits which He foresees that each individual is to have—giving the adoption of sons to those whom He foreknows will not be unworthy of His grace, and dooming those to destruction whose dispositions He perceives will be prone to mischief and wickedness. Thus by interposing foreknowledge as a veil, they not only obscure election, but pretend to give it a different origin. (*Institutes* 3.22.1)

### **3.6 Calvin's Deterministic Convictions**

We can, however, see clearer commitment to a determinist perspective from other things he wrote beyond the *Institutes*. In his *Concerning The Eternal Predestination of God*, which arose out of his dispute with Albertus Pighius (c. 1490–1542) and Georgius of Sicily over Calvin’s doctrine of predestination, it is clear he held to a very intimate view of God’s predestining activity and human action. Here follow pertinent quotes with commentary. I deal first with quotes affirming a belief in the all-determining nature of divine sovereignty. Here is one such broad affirmation:

But of all the things which happen, the first cause is to be understood to be His will, because He so governs the natures created by Him, as to determine all the counsels and the actions of men to the end decreed by Him. (1961 [1552]: 178)

The claim that God is the first cause of all is common in Christendom, and so isn’t readily interpreted as an affirmation of theological determinism. But Calvin’s understanding of what this involves has God governing his creatures’ natures to determine all their ‘counsels and actions’ to the end God has decided. That God can govern human nature to bring about whatever action he desires is exactly the sort of claim a libertarian theist would deny, however. He would say that God cannot generate whatever action God pleases from an individual, because free, morally responsible actions must come from the person themselves—they can’t be caused to perform them. That would make them unfree. But it appears Calvin has no qualms about affirming that God can get whatever actions he wishes from his creatures. This naturally suggests Calvin is a theological determinist and a compatibilist.

Calvin provides a bit more detail when says he sides with Augustine:

Augustine’s opinion is to be accepted: When God wills to be done what cannot be done but by willing men, their hearts being so inclined that they will, He Himself effects this, not only by helping in their hearts but by determining them, so that, though they had no such intention, they fulfil what His hand and His counsel decreed. (1961 [1552]: 176)

The point at issue here is how God, should God wish to bring about a free and willing action from a human being, would bring about that free and willing action. Here Calvin agrees with Augustine that God does not merely bring about desire-states that he believes are likely to tend towards that free action, or that he foreknows will lead to that free action (moves that a libertarian theist would be inclined to make); instead, Calvin affirms that God effects the free action himself, determining men to fulfil what his hand and counsel decree. If Calvin is a libertarian (or proto-libertarian), this is

language that is very unexpected and hard to square with incompatibilist commitments. How can God *effect* the action without what the incompatibilist would take to be a compromising of human freedom?

Moreover, Calvin is clear that these determinations from God can be determinations to either good deeds or evil deeds:

Again it is quite clear from the evidence of Scripture that God works in the hearts of men to incline their wills just as He will, whether to good for his mercy's sake or to evil according their merits (1961 [1552]: 177)

Evidently, Calvin does not think that when God does this determining of someone to perform an evil action, that removes their responsibility—it is still an action for which they are considered guilty and blameworthy. This again indicates a compatibilist understanding.

### **3.7 Calvin's Dual-Actor Principle**

What I consider the most striking evidence of Calvin's commitment to a deterministic perspective is what I term his 'Dual-Actor Principle'. I shall let him speak for himself:

For myself, I take another principle: Whatever things are done wrongly and unjustly by man, these very things are the right and just works of God. (1961 [1552]: 169)

Therefore, every single action, or work, whether good or evil, has two authors: one is man, and the other is God. He elucidates further with reference to the Scriptural example of Job:

Robbers steal the cattle of the saintly Job. The deed is cruel and shameful. Satan by this means tempts him to desperation—an even more detestable machination. But Job himself indicates another author of the deed: The Lord gave, the Lord has taken away. He not unjustly transfers to God what could not be attributed without the robbers. (1961 [1552]: 179–180)

This is very easy to understand on a theological determinist perspective. Suppose God causes a human being to perform a sinful act. The human agent is therefore causally responsible for bringing about the evil intention, but the agent's bringing about that intention is also caused by God. Two agents are therefore causally responsible for the evil intention: the human agent and God. It is true

of both of those agents that they have caused the existence of the evil intention. Of course, the human agent stands in closer causal proximity to the sinful intention—it is *his* intention in a way that it is not God’s—but that does not prevent both actors from being causally responsible for the coming to be of that intention.

It is harder to square Calvin’s Dual-Actor Principle with a libertarian perspective. For on that scheme God does not cause any human agent to act. Or, if he does, then such acts won’t be sinful: because if God *causes* the agent to act, the action cannot be a morally responsible one. In what sense, then, can God be considered the author of human sinful deeds? How can sinful works be justly considered God’s works also? One might say that the wicked acts of sinful men are God’s acts also in the sense that God *arranges* their coming into being. If God foreknows that Peter will sinfully deny Christ in a certain circumstance, and God brings about the circumstance in order that Peter will sin, then the incompatibilist interpretation of Calvin might suggest that that is all Calvin had in mind when he claims that God authors sinful acts.

But this is something of a stretch. It doesn’t appear to license Calvin’s remark that Job ‘not unjustly transfers to God’ what was perpetrated by the robbers. What is true on the libertarian scheme is that God set things up so the robbers could rob, but the robbing itself is the robbers’ own work, and not God’s. Also, if Calvin’s understanding is implicitly libertarian, we would expect to find more of the language of permission from him. We would expect him to say that God arranges matters so that he *permits* a human agent to sin for God’s own purposes, but that no positive ‘shove’ comes from God’s end. The wicked deed is done only by the wicked human agent, and not by God. God has his ‘hands off’, so to speak, when it comes to human sin, even though he may arrange and orchestrate it, and give opportunity for it, for his good purposes. But instead we see Calvin explicitly disclaiming the language of permission:

From this it is easy to conclude how foolish and frail is the support of divine justice afforded by the suggestion that evils come to be not by His will, but merely by his permission. Of course, so far as they are evils, which men perpetrate with their evil mind, as I shall show in greater detail shortly, I admit that they are not pleasing to God. But it is a quite frivolous refuge to say that God permits them, when Scripture shows Him not only willing but the author of them. (1961 [1552]: 176)

Notice that not only does Calvin disclaim the language of permission, he also disclaims the language of mere *willing*, and insists that we must go a step further. In other words, Calvin is concerned to hold that God doesn’t merely plan that men should sin; he wants to say that God does more than that. In fact, he wants to say that Scripture presents God as *author* of the wicked man’s

acts. This is the Dual-Actor principle in action. If a man brings something about, then it is also true that God has brought that thing about. If a man forms a sinful intention in his heart, then it is also true that God has formed that sinful intention in the man's heart. There are always (at least) two authors to any work.

But what could this extra step of authorship amount to? For the determinist interpreter of Calvin, there is no mystery here. Calvin is saying, 'God doesn't merely permit man to sin, nor does he merely will (or plan) that man should sin, he also *determines* man to sin'. That is what is meant by 'authorship' here. But on a libertarian interpretation of Calvin, it is hard to know what the extra thing is that Calvin is insisting on here. On the libertarian view, *all* God can do is plan for sin to occur and to arrange matters so that it does. He cannot cause, or otherwise determine, sin to occur. What, then, could Calvin mean by his further insistence that God 'authors' the sin of the wicked, if he is a libertarian? There doesn't appear to be a satisfactory answer.

### **3.8 Calvin on God as Author of Sin**

Now, to claim that God is the author of the wicked man's acts carries with it an evident problem. How does one block off the natural suspicion that God is therefore guilty of the wicked acts he brings about, and is as wicked as the wicked are? We should hear Calvin's remarks in response to this:

How then is God to be exempted from the blame to which Satan with his instruments is liable? Of course a distinction is made between the deeds of men and their purpose and end; for the cruelty of the man who puts out the eyes of crows or kills a stork is condemned, while the virtue of the judge is praised who puts his hand to the killing of a criminal. (1961 [1552]: 180)

Calvin's response is the natural one to reach for: God's intentions are good; the intentions of the wicked are not. God intends evil for a good end, while the wicked's aim is on his own evil pleasure. This isn't a comprehensive response, but it is a start. Some mention of rights should also be made, I think. God, as sustainer and creator, has the right to bring about evil deeds in his creatures if that suits his purposes, whilst his creatures do not have the right, or not as broadly as God does, to bring about evil intentions (never, I take it, in oneself, and at best rarely in others). But I shall address matters of theodicy in more depth later. I merely intend to place them in one's consciousness here, and to draw attention to Calvin's take.

### 3.9 Calvin on the Ability Requirement for Moral Responsibility

Like Luther, Calvin had remarks (though less bombastic) to make on the ability requirement on moral responsibility so often brought up in objection to the Reformed position. In another dispute Calvin had, this time with Sebastian Castellio (1515–1563), he was pressed on the matter in the following manner by Castellio. Castellio claimed (not unfairly, to my mind) to extract from the writings of Calvin these two claims:

Article 13: From the perspective of God we sin necessarily, whether we sin on account of our own purpose or by accident.

Article 14: Whatever perversions men perpetrate by their own will, those also proceed from the will of God. (from Calvin, 2010 [1558]: 51)

And Castellio objects,

If we sin necessarily, all admonitions are in vain [...] [They are] in vain if it is as impossible for [people] to change as it is for them to swallow a mountain. What if Calvin says that the commands are displayed for the reason that men might be inexcusable? We reply that this is futile. For if you command your son to eat a rock, and he does not do it, he is no more inexcusable after the command than he was before. It is just as if God commands me, “Do not steal,” and then I steal necessarily. I am not more able to abstain from stealing than I am able to eat a rock. I am not more inexcusable after the command than I was before, and I am not more excusable before the command than after it. (from Calvin, 2010 [1558]: 51–52)

He goes on to take a swipe at Calvin’s view of reprobation:

If an impious person is reprobated before he commits impiety, that is, before he is born, from eternity he therefore sins necessarily. He is already inexcusable and condemned before receiving the command from God. This is contrary to all things concerning the law, God, and man. (from Calvin, 2010 [1558]: 52)

We can see what Castellio’s concern is. It’s the Argument from Ability again. His concern in the initial section of the quoted material is with the legitimacy of God’s blaming human beings for

failing to keep his commandments when those human beings' sins are necessitated by the divine will. Why would we not blame a man for failure to swallow a mountain were he under a duty to do so? Because a man is not *able* to swallow a mountain. His inability excuses him from blame. But surely someone who is predestined to sin on some occasion is likewise not able to refrain from sinning on that occasion. Shouldn't his inability excuse him also?

An objection against God's rationality is also present in the quoted material. If man is predestined to sin and unable to refrain from sin, why would God give commands in the first place? To command people to do something they are unable to do is not the mark of a rational agent, or so the thought goes. The response to that, originating in Luther, is that God gives the commands not because he expects their fulfilment, but because he wishes to communicate a sense of guilt or *inexcusability* in the receivers of the command. But Castellio correctly sees that this is not the nub of the issue. The point is that if one cannot be blamed for doing something that is impossible for one to do, then being commanded to do that impossible thing, even commanded by God, cannot increase one's guilt or bring awareness of one's own moral failure, because there was no such moral failure in the first place.

Calvin gets somewhat personal in his response:

I ask you, when last year the hook was in your hand for the purpose of stealing firewood so that you might warm your home, was it not your own will that drove you to steal? If this alone does not suffice for your just condemnation, that knowingly and willingly you disgracefully and wickedly gained at the expense of another, whatever you roar against necessity, this does not in the least secure your acquittal. (2010 [1558]: 112)

Calvin refers to an incident where Castellio was accused of stealing other people's firewood. For his part, Castellio claimed that he was only taking the poorer pieces of driftwood that other people had no claim on.<sup>13</sup>

Putting the firewood to the side, we should note that (i) Calvin does not challenge the idea that people are under a necessity, stemming from God's will, to sin. This is further evidence of Calvin's determinist outlook. We should also note that (ii) Calvin doesn't really engage with the issues Castellio raises about impossibility and inability, and the intuitive pull that some sort of ability condition must be satisfied in order for moral responsibility to obtain. He is content to assert that conscious, informed, voluntary action is sufficient for moral responsibility, ability or no ability. Although this looks like a compatibilist conception of the conditions necessary for moral

---

<sup>13</sup> See Helm (2010: 20, n. 13) for discussion of the affair.

responsibility, and interesting on that account, it is not a comprehensive nor satisfactory response to the intuitions Castellio musters, and in the next chapter I present a more thoroughgoing answer to this class of worry.

### 3.10 Arminius and the Belgic Confession

But it is now time to introduce Jacobus Arminius (1560–1609). Arminius’s deviation from Calvinist thought was occasioned by problems he had with the *Belgic Confession*, a confession first introduced in 1561 and that was popular in the Netherlands and Belgium amongst the Reformed churches of that time.<sup>14</sup> Arminius was a Professor at the University of Leiden, but when his disagreements with the standard understanding of Reformed belief became clear, he fell under the attack of one of his colleagues at the university, Franciscus Gomarus (1563—1641). Gomarus eventually forced Arminius into public disputation on the matter. And in Arminius’s *Declaration of Sentiments*, addressed to the States of Holland and West Friesland in 1608, he summarises the view of predestination he is opposing as follows:

God by an eternal and immutable decree has predestinated, from among men [...] certain individuals to everlasting life, and others to eternal destruction, without any regard whatever to righteousness or sin, to obedience or disobedience, but purely of his own good pleasure, to demonstrate the glory of his justice and mercy; or, (as others assert,) to demonstrate his saving grace, wisdom and free uncontrollable power. (1853: 211–212)

The idea that God elects and reprobates in order to display mercy on the one hand and justice on the other is an idea we will return to. One should also note that Arminius’s expression ‘without any regard whatever to righteousness or sin’ is liable to misinterpretation. The defender of Calvin’s theory of election and reprobation would of course be committed to the idea that God did not elect or reprobate on the *basis* of sin or righteousness, but it is of course true that God cannot satisfy the aim of reprobation unless there is sin. Reprobation is God’s decision to arrange the damnation of an individual, and it remains true that someone cannot be damned unless they sin. Sin is therefore involved as the means by which God implements reprobation. Apart from that, however, Arminius can be said to offer a fair summary.

---

<sup>14</sup> Schaff tells us that ‘The Confession was publicly adopted by a Synod at Antwerp (1566), then at Wesel (1568), more formally by a Synod at Emden (1571) by a national Synod at Dort (1574), another at Middelburg (1581), and again by the great Synod of Dort, April 29, 1619.’ (1931: 505).

He goes on to note an inference from this doctrine:

From this decree of Divine election and reprobation, and from this administration of the means which pertain to the execution of both of them, it follows, that the elect are necessarily saved, it being impossible for them to perish—and that the reprobate are necessarily damned, it being impossible for them to be saved; and all this from the absolute purpose of God, which is altogether antecedent to all things, and to all those causes which are either in things themselves or can possibly result from them. (1853: 215)

Arminius avers that God's decree, on the predestinarian scheme he is opposing, must be the 'antecedent cause' of all things that come to pass, and that God's purpose thereby governs what all entities with causal powers are, both in themselves, and also what they bring about. This is about as close a formulation of theological determinism as one might find in the scholastic period.

But he considers this scheme to be inconsistent with the *Belgic Confession*, and his appeals to the *Belgic Confession* are of direct relevance to understanding Arminius's beliefs on free will and predestination. He makes two arguments, and here is his first:

Without the least contention or caviling, it may very properly be made a question of doubt, whether this doctrine agrees with the Belgic Confession [...]; as I shall briefly demonstrate.

1. In the 14th Article of the Dutch Confession, these expressions occur: "Man knowingly and willingly subjected himself to sin, and, consequently, to death and cursing, while he lent an ear to the deceiving words and impostures of the devil," &c. From this sentence I conclude, that man did not sin on account of any necessity through a preceding decree of Predestination: which inference is diametrically opposed to that doctrine of Predestination against which I now contend. (1853: 220)

We can see at once what Arminius's issue was: he was an *incompatibilist*. If man knowingly and willingly fell when he listened to the overtures of the devil, then it follows that he couldn't have been caused or determined or necessitated to do so. Because then it wouldn't have been free! And man couldn't be blamed for it. Arminius reads the Belgic Confession in the light of his intuitive incompatibilism, and thereby perceives a tension with the predestinarian doctrine in circulation.

He is explicit further on in the *Sentiments* about his incompatibilism:

This [predestinarian] doctrine is inconsistent with the freedom of the will, in which and with which man was created by God. For it prevents the exercise of this liberty, by binding or determining the will absolutely to one object, that is, to do this thing precisely, or to do that. (1853: 224)

And also:

This Predestination is inconsistent with the Nature and Properties of Sin [...] [b]ecause sin is called “disobedience” and “rebellion,” neither of which terms can possibly apply to any person who by a preceding divine decree is placed under an unavoidable necessity of sinning. (1853: 227)

That is Arminius’s first objection on the basis of the *Belgic Confession*. He has a second:

Then, in the 16th Article, which treats of the eternal election of God, these words are contained: “God shewed himself Merciful, by delivering from damnation, and by saving, those persons whom, in his eternal and immutable counsel and cording to his gratuitous goodness, he chose in Christ Jesus our Lord, without any regard to their works. And he shewed himself just, in leaving others in that their fall and perdition into which they had precipitated themselves.” It is not obvious to me, how these words are consistent with this doctrine of Predestination. (1853: 220–221)

It isn’t entirely plain what Arminius’s problem is here. I think he must be taking issue with either (i) the thought that reprobates have ‘precipitated themselves’, or that (ii) that God ‘shewed himself just’ in the act of reprobation. If the first, then his point amounts to nothing more than another expression of his incompatibilism: how could they be justly said to precipitate themselves when they fall because God has predestined them to fall? If the second, however, then a different point is being made. If Arminius is complaining about the idea that God could show himself *just* by reprobating, then he must think that it would be in tension, somehow, with God’s justice for him to reprobate. But what is the argument there? Later on in his *Sentiments*, we find a candidate for his argument:

[S]in is the meritorious cause of damnation. But the meritorious cause which moves the Divine will to reprobate, is according to justice; and it induces God, who holds sin in abhorrence, to will reprobation. Sin, therefore, which is a cause, cannot be placed among the means, by which God executes the decree or will of reprobation. (1853: 227)

So, a plausible reconstruction of the argument is this: A just God abhors sin; but the end of reprobation is damnation, and damnation requires the existence of sin; therefore, reprobation requires God to will sin, and that is contrary to the just nature of God. He makes similar remarks when he says that reprobation is opposed to the goodness of God:

[Reprobation] is also repugnant to the Goodness of God. Goodness is an affection [or disposition] in God to communicate his own good so far as his justice considers and admits to be fitting and proper. But in this doctrine the following act is attributed to God, that, of himself, and induced to it by nothing external, he wills the greatest evil to his creatures; and that from all eternity he has pre-ordained that evil for them, or pre-determined to impart it to them, even before he resolved to bestow upon them any portion of good. For this doctrine states, that God willed to damn; and, that he might be able to do this, be willed to create; although creation is the first egress [...] of God's goodness towards his creatures. (1853: 223)

The objection here is slightly different, however. In this latter case the objection appears to centre on God's goodness as a disposition to do good. If God is disposed to do good to his creatures, then why is he so disposed to as to will and arrange such a terrible end for so many of his creatures? Reprobation is therefore in tension with God's goodness.

Lastly, Arminius also complains that

Reprobation is an act of hatred, and from hatred derives its origin. But creation does not proceed from hatred; it is not therefore a way or means, which belongs to the execution of the decree of reprobation. (1853: 225)

Whether or not Arminius is right that creation does not proceed from hatred, there is a problem over the bare fact that reprobation involves, or appears to involve, hatred. How could God hate an individual before they are even created? Before they have offended him? And isn't God a God of love? How, then, is any expression of hatred consistent with God's loving nature? These questions are also pressed by contemporary critics of Calvinism, and we shall come to them in due course.

### **3.11 The Remonstrants and the Synod of Dort (1618–1619)**

With Arminius's argumentation, the great split in early Protestant thought was effected and it has persisted to this day. At the present time, evangelical Christianity is divided between the Arminians and the Calvinists. Now, Arminius died in 1609, and was therefore unable to further publicly defend his creed. But his followers took up the slack, and in 1610 published the *Five Articles of Remonstrance*. (Their party then became known as the 'Remonstrants'.) The *Five Articles* do not

explicitly repudiate the strong predestinarian doctrine of Calvin, however, save, perhaps, in the first and fourth article. The first article contains this:

That God, by an eternal, unchangeable purpose in Jesus Christ his Son, before the foundation of the world, hath determined, out of the fallen, sinful race of men, to save in Christ, [...] those who, through the grace of the Holy Ghost, shall believe on this his Son Jesus (Schaff, 1977: 545)

The force of this phrasing is that it avoids commitment to the idea that God has predestined *individuals* to either salvation or damnation. The article is readily interpreted as meaning that the object of God's election was a certain *description*, namely, 'those who shall believe on his Son Jesus'. Thus, God elects a certain description as being one that saves if one satisfies it, and leaves it to man's own free will to meet that description. In this way, the article can be seen as dodging Calvin's commitment to the unconditional election and reprobation of individuals.

The fourth article contains this:

But as respects the mode of the operation of this grace [the grace whereby man accomplishes something good], it is not irresistible, inasmuch as it is written concerning many, that they have resisted the Holy Ghost. (Schaff, 1977: 547)

An anti-determinist concern might be perceived here, insofar as 'irresistible grace' can be understood as 'determining grace'. If so, then the concern of this article is to affirm that God's grace doesn't cause, or otherwise determine, one to a good action (such as placing one's faith in Christ); instead, whether the action is performed or not is presumably left to one's own (libertarian) free will. Thus, it is easy to see incompatibilist convictions at play in these early Arminian declarations.

The Synod of Dort (or Dordrecht) was held from November 1618 to May 1619 in order to address the Remonstrant doctrine and the civil conflict it was causing. However, the *Canons of Dort*, which were the final and considered declarations of the synod on the issue, refused to affirm the stronger view of predestination—that all things are necessitated, or determined, by God's decree. They did say, in Rejection VII of the 'Third and Fourth Main Points of Doctrine', that they condemned those

Who teach: That God in the regeneration of man does not use such powers of His omnipotence as potently and infallibly bend mans will to faith and conversion (Anon., 2010: 24)

This ruled out Arminius's view that man's coming freely to Christ was to be understood in libertarian terms. But the *Canons* also contained material that was not altogether friendly to the deterministic perspective, even though it didn't explicitly rule it out. For instance, in Article VII of under the 'First Head of Doctrine', the *Canons* declare:

Election is the unchangeable purpose of God, whereby, before the foundation of the world, He hath out of mere grace, according to the sovereign good pleasure of His own will, chosen, from the whole human race, which had fallen through their own fault from their primitive state of rectitude into sin and destruction, a certain number of persons to redemption in Christ (Anon., 2010: 9)

The *Canons*'s affirming that divine election takes place after, in the order of God's decisions, man has 'fallen through his own fault',<sup>15</sup> carries the natural implication that humanity's fall was not something that was necessitated, or predestined, or ordained by God. Indeed, there is no pronouncement that everything that comes to pass is determined by God's decree—the *Canons* restricts its discussion to matters of salvation only. The *Canons* are therefore best read as an affirmation of soteriological Calvinism rather than determinist Calvinism.

Even worse from the deterministic perspective, the *Canons* also declare, in Article 5 under the 'First Head of Doctrine', that

The cause or guilt of this unbelief [in Christ], as well as of all other sins, is no wise in God, but in man himself (Anon., 2010: 9)

This doesn't constitute a denial that God is the ultimate cause of man's sin (and therefore a denial of theological determinism), because 'cause' appears to be used there as synonym (or close to that) of 'guilt'. With regard to guilt, the statement is to be understood as declaring that the guilt-worthiness of the act is grounded in features of man, not features of God. And that is surely correct: what makes an act a wicked act—the evil intentions—these are present only in man, and never in God. 'Cause' should therefore be understood in a similar way, to mean something like 'active principle', or 'reason', or 'ground'. And indeed the wicked principles of man's nature that lead to unbelief and the sinful features of his psychology in which unbelief is grounded, are again to be found only in man, and not in God.

---

15 The *Canons*' claim here is only properly understood within the context of the debate between infralapsarians and supralapsarians. The former are seen as offering a softer view of predestination than the latter. Infralapsarians claim that God only decided to elect and reprobate after he had decided (for some other reason) to bring about the fall. God thus worked with human 'already fallen'. The supralapsarians claimed that God's decrees of election and reprobation preceded the fall. Thus, God was electing and reprobating 'before' humanity was fallen.

But even if there is nothing to strictly rule out theological determinism in the *Canons of Dort*, there is little encouragement to be found for theological determinists in those *Canons*. We will have to wait until we reach the celebrated *1646 Westminster Confession* before we see the contours of Reformed thought return to reasonably clear affirmations of theological determinism.

### 3.12 Molinism

But before we discuss the Westminster Confession, another signal development in philosophical theology must be mentioned, that of Molinism. ‘Molinism’ is the term used to refer to the theory of Luis de Molina (1535–1600), a Roman Catholic theologian. For the debate in Reformed thought between the Calvinists and the Arminians was paralleled in the Catholic tradition by the debate between the Dominican Thomists and the Jesuit Molinists. In the De Auxiliis Controversy (controversy on help), the two parties debated the nature of the help that God gave to man when God bestows saving grace upon him.<sup>16</sup> Both parties shared the Catholic commitment that the grace by which God brings about an individual’s salvation is of infallible efficaciousness in bringing about the consent of the saved party. But the two parties disagreed on the mechanism by which this was accomplished. The Thomists thought that God secured consent through ‘a physical impulsion by means of which God determines and applies our faculties to the action’ (Astrain 1908). The Molinists thought that the efficacy of the grace followed from God’s *middle knowledge*, a knowledge of what any individual would freely choose to do in any possible circumstance. By only giving saving grace to those who God foreknows will freely choose him, God can thereby guarantee that the individual offered grace will always accept it. Thus, infallible grace.

Molina laid out his theory in his *Concordia* (1988 [1588]), wherein he distinguished between three types of knowledge in God. The first was God’s natural knowledge. This is God’s knowledge of all necessary truth (such as mathematical truths). This knowledge is had by God prior to God’s decision to create. Indeed, this is part of the knowledge that God relied on when deciding to create and deciding what to create, because this knowledge contains all the necessarily true subjunctive conditionals, truths like ‘were I to decide the laws of nature are going to be *this* way, then events of such-and-such a sort would be nomologically impossible’. God surely decides what to create in the light of such truths. Also note that this natural knowledge is *essential* to God. God can’t be God without it. There can’t be a possible world where God doesn’t have natural knowledge because

---

<sup>16</sup> See Matava 2016 for an outline of the controversy.

there is no possible world where the propositions which are naturally known are false, and God knows everything true in every possible world in which he exists (which is all possible worlds).

The next sort of knowledge is free knowledge. This is God's knowledge of contingent matters of fact that obtain as the result of God's creative decree. Suppose God decides to create Adam. With Adam in existence, God would know the proposition <Adam exists> to be true. But this knowledge is not essential to God. God, on the standard understanding, could have refrained from creating Adam. God did not create Adam necessarily. Therefore, there is a possible world in which Adam does not exist. In that world, God does not know the proposition <Adam exists> to be true; he knows it to be false. Free knowledge is inessential to God and it is God's knowledge of things that God brings about by his free choice.

Summing up, God's natural knowledge is (i) prior to his decision to create and (ii) essential to God (because knowledge of things necessary). God's free knowledge is (iii) posterior to creation and (iv) not essential to God (because knowledge of things contingent). Molina's suggestion is that there is a third category of knowledge between these two: *middle knowledge*. Middle knowledge holds (v) prior to God's decision to create, but it is (vi) inessential to God (because it is knowledge of contingent matters).

What sort of things are the objects of God's middle knowledge? Molinists believe that all propositions about what human beings would freely choose to do in any possible circumstance are included as part of the objects of middle knowledge. 'Subjunctive conditionals of freedom', as they are known. These are propositions of the following form:

Were an agent, *S*, to be placed in circumstances, *C*, *S* would freely *X*,

where *X* ranges over possible decisions. Thus, Molinism posits that, before God can form any intention to create anything, God is aware of an infinity of truths about what any possible person would freely choose to do in any possible circumstance. And no particular truth of that kind will be essential to God. Suppose, as a matter of fact, that it is true that

(1) Were Curley to be offered \$100, he would freely accept the bribe.

For the Molinist, that truth cannot be necessarily true. Molinists are libertarians. If it is necessarily true that Curley would accept the bribe, then it can't be a *free* accepting of the bribe. It must therefore be a contingent truth, and if it is contingent, then it can't be essential to God's nature to know it. These are all therefore contingent truths about what any possible person would freely do in

any possible circumstance, and they are all thrust into God's mind, alongside his natural knowledge, before God can decide to create.

That is how Molinists understand God's working providentially: God foreknows what each of his creatures would freely do in any situation, and God therefore brings about the situations that God foreknows will lead to the free decisions God wants to bring about. And that is how Molinists solve the *de auxiliis* problem of how it is that grace could infallibly bring about the free consent of a human person: God uses middle knowledge to bring about those circumstances he foreknows would lead to the free compliance of the agent.

Why introduce Molinism to the discussion? Several reasons. For one thing, Arminius read Molina. He had a copy of the *Concordia*, and Arminius's language is considered to reflect Molina's.<sup>17</sup> More generally, it is being increasingly realised that this early debate between Calvinists and Arminians at the 1618–19 synod of Dort owed a great deal to the Roman Catholic discussions of grace and free will that took place during the *de auxiliis* controversy that ran just a few decades before, from 1581 to 1607.<sup>18</sup> And, as we shall see, it will go on to affect the statement of divine sovereignty we find in the *1646 Westminster Confession*. For another thing, it is the most popular formulation of Arminianism current today. William Lane Craig, a popular contemporary defender of Molinism, writes that it is 'one of the most fruitful theological ideas ever conceived. For it would serve to explain not only God's knowledge of the future, but divine providence and predestination as well' (2000: 127). The Molinist theory of divine providence remains the chief rival to the theological determinist one. On that score alone, it is good to bear it in mind. Lastly, the theory will have an effect on certain other discussions in this thesis.

### 3.13 The 1646 Westminster Confession

Finally, we come to the *Westminster Confession of Faith* of 1646. This confession, with the associated *Shorter* and *Larger Catechisms*, produced by the Westminster Assembly of Divines (which met 1643–53) are widely regarded as the greatest expression of the Calvinistic, Reformed faith. Fesko writes that 'The Confession and catechisms of the Westminster Assembly have been praised by theologians, both in the seventeenth century and in our own day, as being the high-water mark of Reformed theology in the early modern period' (2014: 23). Schaff writes of the Westminster Assembly,

---

17 Dekker (1996) 'was Arminius a Molinist?'

18 See the soon-to-be-released volume edited by Ballor, Gaetano, and Sytsma (forthcoming).

Whether we look at the extent or ability of the labors, or its influence upon future generations, it stands first among Protestant Councils. The Synod of Dort was indeed fully equal to it in learning and moral weight, [...] but the doctrinal legislation of the Synod of Dort was confined to the five points at issue between Calvinism and Arminianism; the Assembly of Westminster embraced the whole field of theology, from the eternal decrees of God to the final judgement. The Canons of Dort have lost their hold upon the mother country; the Confessions and Shorter Catechism of Westminster are as much used now in Anglo-Presbyterian Churches as ever, and have more vitality and influence than any other Calvinistic Confession. (1931: 728)

If this great expression of Reformed convictions reproduces the deterministic strain present in Calvinist thought, then that is a great feather in the cap of the theological determinist who seeks to ground his beliefs, to a substantial extent, in the Reformed tradition.

I shall note three ways in which the *Westminster Confession* moves beyond mere soteriological Calvinism—in his fallen state, man is bound, by necessity to sin—into the commitments of deterministic or philosophical Calvinism—the idea that God’s decree determines (or, in their language, necessitates, all that comes to pass).

### **3.14 The Determinism of the Westminster Confession: (i) God’s Decree**

Here is the *Westminster Confession* ch. III, on God’s Eternal Decree, 1–2:

God from all eternity did by the most wise and holy counsel of his own will, freely and unchangeably ordain whatsoever comes to pass; [...] Although God knows whatsoever may or can come to pass, upon all supposed conditions; yet hath he not decreed any thing because he foresaw it as future, or as that which would come to pass, upon such conditions. (1977: 608)

First note that the *Confession* insists that God ordains *all* that comes to pass. That includes the free actions of men. Of course, ‘ordain’ may be given an incompatibilist interpretation. But what follows tells against that idea. An incompatibilist understanding of what it is to ordain a free action of the creature would have it as the Molinist has it: God decrees based on foreknown free action. But that is precisely what the Westminster Assembly is concerned to rule out. For here we can see how the Westminster divines responded to Molinist ideas. They categorically deny that God used middle knowledge—knowledge of what agents would freely do under certain conditions—when deciding

what to create. God decrees *nothing*, in their eyes, on the basis of events either foreknown simply, or foreknown to come to pass under certain conditions. The Arminian, incompatibilist understanding is therefore ruled out.

Here is an extract from ch. V of the *Confession*, ‘Of Providence’, section 1:

God, the great Creator of all things, doth uphold, direct, dispose, and govern all creatures, actions, and things, from the greatest even to the least, by his most wise and holy providence, according to his infallible foreknowledge (1977: 612)

It is to be noted that God’s directing and disposing of all things explicitly includes human *actions*. It is not impossible, perhaps, to reconcile a statement like that with a libertarian understanding, but an insistence that God directs and disposes every human action offers the libertarian little comfort.

This portion from section 4 of the same chapter tells more strongly against both libertarianism and mere soteriological Calvinism:

The almighty power, unsearchable wisdom, and infinite goodness of God, so far manifest themselves in his providence, that it extendeth itself even to the first Fall, and all other sins of angels and men, and that not by a bare permission, but such as hath joined with it a most wise and powerful bounding, and otherwise ordering and governing of them, in a manifold dispensation, to his own holy ends; yet so, as the sinfulness thereof proceedeth only from the creature, and not from God; who being most holy and righteous, neither is nor can be the author or approver of sin. (1977: 613)

Of chief interest is the expression ‘and that *not by a bare permission*’. The *Confession* insists that God doesn’t take his hands off with human action, even sinful human action, but joins ‘with it a most wise and powerful bounding’, a bounding which guarantees that the action takes place. For my part, I cannot see what this bounding could be save for a determining relation. I therefore cannot see how the libertarian, because of his insistence that morally responsible action requires indeterminism, can agree with this paragraph from the *Confession*. The libertarian picture is precisely the picture of bare permission—God actualises the relevant circumstance, *a la* Molinism, and then watches man’s actions unfold as he foreknows they will. But God cannot determine the action itself, if libertarianism is true. He must merely permit it. But the Westminster divines insist there is more going on than mere permission. Insofar as the Westminster Assembly speaks for the Reformed tradition, this constitutes a very strong piece of evidence that the Reformed tradition is deterministic.

Not only that, this section also tells against mere soteriological Calvinism. That, if you recall, is the sort of Calvinism that insists that postlapsarian humanity (man after the fall) is bound by necessity to sin, but man before the fall, and man after regenerating grace, may well be able to perform acts that are free in the incompatibilist sense. This section tells against that view because it includes in its remit Adam's sin. It says that divine providence, this 'bounding', 'extendeth itself even to the first Fall, and all other sins of angels and men'. Thus, Adam's eating of the fruit, an action performed before the canker of sin got a hold of man's nature, was also infallibly bounded to its occurrence by the power of God. In this way, the *Westminster Confession* corrects and enlarges the weaker and more restricted posture displayed in the *Canons of Dort*.

### **3.15 The Determinism of the Westminster Confession: (ii) 'Coming Most Freely'**

The second way the *Westminster Confession* gives support to a deterministic perspective is through the expression 'come most freely' in Chapter X, 'Of Effectual Calling'. Speaking of how it is that God calls a soul—that is to say, regenerating it from its native darkness and drawing it to Christ—the *Confession* says that God:

enlighten[s] their minds, spiritually and savingly, to understand the things of God, taking away their heart of stone, and giving unto them an heart of flesh; renewing their wills, and by his almighty power determining them to that which is good; and effectually drawing them to Jesus Christ; yet so as they come most freely, being made willing by his grace. (1977: 624)

The *Confession* claims both that such drawing is 'effectual' and that that effectuality is owed to the exercising of God's 'almighty power', an exercising that 'determines them to what is good'. If God is effectually exercising his power and thereby determining his elect to the good, that looks like a clear case of determining of the stricter, philosophical sort that this thesis is concerned with. That by itself might be not of great significance, perhaps, but the fact that the Assembly then go on to aver that, nevertheless, the effectually called individual 'comes *most freely*' cannot escape notice. It appears that, even though they wouldn't, and couldn't, have said so explicitly, the Westminster divines are compatibilists. They think that it is possible for God to determine an agent to come to Christ, yet for it still to be true that that agent comes freely.<sup>19</sup> But if the *Westminster Confession* entails compatibilism, then it undermines pressure to deny theological determinism, because the

---

<sup>19</sup> The same point is made by Anderson and Manata (2017: 294–295).

denial of theological determinism is nearly always motivated by the desire to accommodate incompatibilistic convictions.

### 3.16 The Determinism of the Westminster Confession: (iii) Divine Aseity

The third way in which the *Westminster Confession* implies theological determinism is through its doctrine of aseity. In Chapter II, 'Of God, and of the Holy Trinity', section 2, we read this:

God hath all life, glory, goodness, blessedness, in and of himself; and is alone in and unto himself all-sufficient, not standing in need of any creatures which he hath made [...] In his sight all things are open and manifest; his knowledge is infinite, infallible, and independent upon the creature; so as nothing is to him contingent or uncertain. (1977: 607)

Here we are informed that God is 'all-sufficient' and that his knowledge is 'independent upon the creature', that is to say, for anything God knows about his creatures, that knowledge is not dependent upon the creature. If God's being omniscient (his knowing everything) depended upon the creature, then God would not be *a se*, all-sufficient, in the sense the Westminster divines believe he is. How, then, *does* God know anything about the creature? If God's knowledge is not based on the reality of the creature, what is it based on? The answer is that God's knowledge of creation is based on his decree. God knows that Albert is six-feet tall because God has decreed that Albert should be 6-feet tall, not because of Albert's being six-feet tall. One corollary of this model is that everything external to God must be decreed by God, else God could not know of it. But the Arminian will claim that either the free actions of human beings are independent of God, or, *a la* Molinism, facts about what agents would freely do are independent of God. Either way, there will be something that God's decree does not determine, and therefore something that God cannot know, given the Westminster Assembly's understanding of aseity. The *Westminster Confession* is therefore committed to the impossibility of divine middle knowledge. Anderson and Manata summarise the argument:

God *alone* is the source of his eternal decree. God doesn't "consult" anything *extra se* when he formulates his decree. To put the point in a quasi-syllogistic form: every event takes place according to God's eternal decree; God's eternal decree is not determined in any respect by anything external to or independent of God; therefore, every event is ultimately determined by God alone. (2017: 286)

### 3.17 The Westminster Confession Opposed to Determinism?

I think all these three preceding considerations lend solid weight to the claim that the Reformed tradition, as encapsulated by the Westminster Assembly, is deterministic. This goes against the views of some commentators. Cunningham, for instance, states that

there is nothing in the Calvinistic system of theology or in the Westminster Confession which requires men to hold the doctrine of philosophical necessity; or in other words, that a man may conscientiously assent to the Westminster Confession although he should reject that doctrine. (1862: 508)

But, as we have seen, that does not agree with a close reading of the *Confession*.

Now, there are some portions of the *Confession* which appear to give support to a libertarian perspective, and these have been capitalised on by opponents of theological determinism who identify as Reformed. J. V. Fesko, for example, believes that in ch. III, paragraph 1 of the *Confession* the Westminster divines explicitly ruled out a deterministic understanding of divine providence. The relevant passage runs as follows:

God from all eternity did by the most wise and holy counsel of his own will, freely and unchangeably ordain whatsoever comes to pass; yet so as thereby neither is God the author of sin; nor is violence offered to the will of the creatures, nor is the liberty or contingency of second causes taken away, but rather established. (1977: 608)

(I quoted this section earlier, but took out certain portions. Here it is quoted fully.) Matters therefore seem plain to Fesko, who remarks:

the Confession does not teach philosophical determinism (or necessity) and does affirm contingency (2014: 99)

He elaborates,

Contingency does not mean that something does not have a cause [...] Rather, it means that something could be otherwise. [...] But once God decrees it, there is no longer contingency from the

divine perspective [...] As it pertains to creatures, however, the divines state that the decree, far from taking away freedom and contingency, establishes it. (2014: 103)

But it is confused to think that these remarks in the *Confession* on contingency express an aversion to philosophical determinism. Then what did the *Confession* mean by its insistence on contingency? To answer that we will need to understand with the way that mediaevals approached this issue—and it is somewhat strange to the modern mind. If you suggested to a Reformed scholastic that God's decree necessitates its objects, and therefore that if God decrees that S will choose to A, then S's choosing to A is necessary, and on that account not free (because freedom and necessity are incompatible), he would respond by saying that you commit a modal fallacy. Let 'N' be the necessity operator. The scholastic would say it does not follow from

(2)  $N(p \rightarrow q)$

that

(3)  $N(q)$

It can therefore be true both that

(4)  $N(\text{God decrees that } S \text{ chooses to } \varphi \rightarrow S \text{ chooses to } \varphi)$

and that

(5)  $\sim N(S \text{ chooses to } \varphi)$ .

In short, one is not entitled to infer the necessity of the consequent from the necessity of the consequence. One can therefore cheerfully maintain the contingency of free decisions even in the light of God's necessarily successful decree.<sup>20</sup>

But the problem is that the majority of modern-day advocates of soft compatibilism would happily grant that free decisions are contingent and not necessary in the way (5) expresses. Because determinism doesn't entail anything like that. Determinism is in the business of the necessity of the consequence, not the consequent. It says that some set of facts determines another set of facts; it

---

<sup>20</sup> See van Asselt, Bac & te Velde (2010: 35–38) for a good summary of the scholastic understanding of the issue.

doesn't say that that latter determined set of facts is necessarily determined—in modern parlance, determined in every possible world. Perhaps the determined set of facts is sometimes (across the space of possibilities) determined, and sometimes not. Perhaps the nature of the determined facts is equally capable of being both determined and undetermined, even though they are, in the actual case, determined.

So, given that the determinist can grant what the *Confession* says at this point, it can't be insisted that the *Confession* rules out theological determinism. There is therefore no real pushback against the earlier three arguments that the theology of the Westminster Assembly favoured what is fundamentally a determinist outlook.

### 3.18 Wesley, Channing, and Westminster

There are two remaining characters I wish to introduce: John Wesley and William Channing. They are both known for their strident moral objections to Calvinism. In particular, they see the Calvinist doctrines of election and reprobation as casting grave doubt on the goodness of God. I want to present their key arguments in this section.

The *Westminster Confession* reaffirmed Calvin's account of election and reprobation as unconditional. On election the *Confession* (III: 5) says:

Those of mankind that are predestinated unto life, God, before the foundation of the world was laid, according to his eternal and immutable purpose, and the secret counsel and good pleasure of his will, hath chosen in Christ, unto everlasting glory, out of his free grace and love alone, without any foresight of faith or good works, or perseverance in either of them, or any other thing in the creature, as conditions, or causes moving him thereunto; and all to the praise of his glorious grace. (1977: 609)

And on reprobation the *Confession* (III: 7) likewise says:

The rest of mankind, God was pleased, according to the unsearchable counsel of his own will, whereby he extendeth or withholdeth mercy as he pleaseth, for the glory of his sovereign power over his creatures, to pass by, and to ordain them to dishonour and wrath for their sin, to the praise of his glorious justice. (1977: 610)

It is this decision of God to ‘pass by’ the non-elect, and thereby guarantee their damnation, when it appears he could have easily made every human being one of his elect, that particularly raises the ire of Wesley and Channing.

### 3.19 John Wesley

John Wesley (1703–91) was one of the founders of the Methodist movement, and a very influential figure in the history of 18<sup>th</sup> century England. The Methodist movement was a breakaway from the established Anglican church and its attendant strictures, and the movement was composed of Calvinist and Arminian factions. George Whitefield (1714–1770) is fairly considered the leader of the Calvinist contingent, though Wesley’s relationship with Whitefield was friendly. He did, however, have a fierce exchange over predestination and the like with Augustus Toplady (1740–1778), and Wesley’s attacks on Calvinism are perhaps best represented in his work *Predestination Calmly Considered*, and his (not so calm) sermon of 1740 entitled ‘Free Grace’.

In the former work he states his convictions on the matter of an unconditional decree of reprobation:

[U]nconditional election I cannot believe; not only because I cannot find it in Scripture, but also (to wave all other considerations) because it necessarily implies unconditional reprobation. Find out any election which does not imply reprobation, and I will gladly agree to it. But reprobation I can never agree to while I believe the Scripture to be of God; as being utterly irreconcilable to the whole scope and tenor both of the Old and New Testament. (1997: 250)

But what, exactly, is the nature of his complaint? One issue is one that, as we have seen, is raised so often in critiques of Calvinism: the issue of ability and human responsibility. Wesley thought that unconditional reprobation (and predestination more generally) is incompatible with God’s holding man responsible:

The sovereignty of God is then never to be brought to supersede his justice. And this is the present objection against unconditional reprobation; (the plain consequence of unconditional election;) it flatly contradicts, indeed utterly overthrows, the Scripture account of the justice of God. [...] The Scripture describes God as the Judge of the earth. But how shall God in justice judge the world? (O consider this, as in the presence of God, with reverence and godly fear!) How shall God in justice judge the world, if there be any decree of reprobation? On this supposition, what should those on the

left hand be condemned for? For their having done evil? They could not help it. There never was a time when they could have helped it. God, you say, “of old ordained them to this condemnation.” And “who hath resisted his will?” He “sold” them, you say, “to work wickedness,” even from their mother’s womb. He “gave them up to a reprobate mind,” or ever they hung upon their mother’s breast. Shall he then condemn them for what they could not help? (1997: 261–262)

By now a familiar argument. If God has predestined, and let us also say, determined, an individual to perform some wicked act, such as rejecting Christ, then how can God blame them for thus acting wickedly? After all, a determining decree from God appears to give them the perfect excuse: they couldn’t help it! If God has decreed it, then it must come to pass. No one can change that. How, therefore, can it be that one is blamed for what one cannot change? The reprobate lacks ability; therefore, he lacks moral responsibility. So goes the argument.

Wesley then proceeds to make a different claim. He claims that reprobation is inconsistent with the *love of God*:

So ill do election and reprobation agree with the truth and sincerity of God! But do they not agree least of all with the scriptural account of his love and goodness? that attribute which God peculiarly claims, wherein he glories above all the rest. It is not written, “God is justice,” or “God is truth:” (Although he is just and true in all his ways:) But it is written, “God is love,” love in the abstract, without bounds; and “there is no end of his goodness.” His love extends even to those who neither love nor fear him. He is good, even to the evil and the unthankful; yea, without any exposition or limitation, to all the children of men. For “the Lord is loving” (or good) “to every man, and his mercy is over all his works.”

But how is God good or loving to a reprobate, or one that is not elected? (1997: 268)

This is a different complaint, but again a fairly natural one. If God is love, and God is loving to all, then isn’t that inconsistent, or at least in tension, with an unconditional decree of reprobation? If God is loving to all wouldn’t that lead him to electing all? Moreover, if God loves all, then he must love the reprobate. But how can it in good conscience be said that God loves one he has predestined to eternal damnation?

In his sermon, ‘Free Grace’, Wesley’s rhetoric is much more strident. He says that Calvinism

destroys all [God’s] Attributes at once. It overturns both his Justice, Mercy and Truth. Yea, it represents the most holy GOD as worse than the Devil; as both more false, more cruel and more unjust. More False; because the Devil, Liar as he is, hath never said, He willeth all Men to be saved.

More Unjust, because the Devil cannot, if he would, be guilty of such Injustice as you ascribe to GOD, when you say, That GOD condemned Millions of Souls to everlasting Fire prepared for the Devil and his Angels for continuing in Sin, which for want of that Grace he will not give them, they cannot avoid; And more Cruel, because that unhappy Spirit seeketh Rest and findeth none; so that his own restless Misery is a kind of Temptation to him to tempt others. But GOD resteth in his high and Holy Place: So that to suppose him of his own mere Motion, of his pure Will and Pleasure, happy as he is, to doom his Creatures, whether they will or no, to endless Misery; is to impute such Cruelty to him, as we cannot impute even to the great Enemy of GOD and Man. It is to represent the most High GOD (He that hath Ears to hear, let him hear!) as more Cruel, False, and Unjust than the Devil. (1741: 24–25)

Three claims are made here: that Calvinism makes God (i) false, (ii) unjust, and (iii) cruel. The claim that Calvinism makes God false appears to come from verses like 1 Timothy 2:3–4, which read, ‘God our Saviour, who desires all people to be saved and to come to the knowledge of the truth.’ The objection must be that, if God has reprobated certain people, then it can’t be said that he desires their salvation. Otherwise, wouldn’t he have elected them? God is thought to be more unjust because of the ability-based reasons given previously: if they can’t avoid God’s decree for them, how can they be blamed? God is held to be cruel because there appears to be no *need* for God to reprobate. What could motivate God to do such a thing?

And yet more objections remain. William Channing is an able mouthpiece for them.

### **3.20 William Channing**

William Channing (1780–1842) was an influential Unitarian theologian and preacher of New England. But early on in his intellectual development he came to oppose the Calvinist doctrine he encountered. In 1820 he published a tract entitled ‘The Moral Argument Against Calvinism’. The arguments found there are cut from the same cloth as Wesley’s, and complement them well. Like Wesley, Channing is naturally repelled by the position of the Westminster Assembly:

Whoever will consult the famous Assembly’s Catechisms and Confession, will see the peculiarities of the system in all their length and breadth of deformity. A man of plain sense, whose spirit has not been broken to this creed by education or terror, will think that it is not necessary for us to travel to heathen countries, to learn how mournfully the human mind may misrepresent the Deity. (1841: 223)

But what, exactly, is the difficulty with the *Confession* and the catechisms? Channing says:

[T]he principal argument against Calvinism, in the General View of Christian Doctrines, is the moral argument, or that which is drawn from the inconsistency of the system with the divine perfections. It is plain, that a doctrine, which contradicts our best ideas of goodness and justice, cannot come from the just and good God, or be a true representation of his character. (1841: 221–222)

He goes on:

Christianity, we all agree, is designed to manifest God as perfect benevolence, and to bring men to love and imitate him. Now is it probable, that a religion, having this object, gives views of the Supreme Being, from which our moral convictions and benevolent sentiments shrink with horror, and which, if made our pattern, would convert us into monsters! It is plain, that, were a human parent to form himself on the universal Father, as described by Calvinism, that is, were he to bring his children into life totally depraved, and then to pursue them with endless punishment, we should charge him with a cruelty not surpassed in the annals of the world (1841: 238)

The complaint that Calvinism calls into question God's benevolence can be fairly said to be covered under the already mentioned argument the Calvinism undermines the love of God, but this section is to be remarked on because it raises the issue of reprobation in the light of the *fatherhood of God*. A complaint like that goes beyond concerns about mere divine benevolence. If it could be shown that reprobation is compatible with divine benevolence, the question would remain whether it is compatible with God as *father* of humanity. As Channing points out, if an earthly father were to reprobate his offspring, then he would be guilty, intuitively, of a great failure in paternal duty. Doesn't Calvinism imply that God is likewise guilty?

We also appear to see in Channing here a dislike of Hell, endless punishment, itself. He appears to think that that doctrine too, which is very much a part of the Reformed confessions, also impugns the goodness of God by attributing to him a cruelty that no earthly ruler could equal. Compared to the endless suffering of the damned, any torment occurring on this side of the grave will inevitably come up short. There is a real worry here about the justice of Hell: how can it be just for God to punish man with an infinite, or endless, punishment for his sins, which appear only finite?

It is the task of this thesis to respond to all these objections, and the thesis is structured around successive responses to these objections against Calvinism collated in this chapter.

### 3:21 Summing Up

This chapter has been somewhat convoluted, so it is worth summarising the main points. We began by examining Luther's position on free will, and, despite tensions in his thought, we saw material there that suggested Luther's perspective was at least sympathetic to theological determinism. When we looked at Calvin's writings, we encountered a far plainer commitment to deterministic thinking. The Synod of Dort, which convened in response to Arminian theology, was disappointing in its refusal to endorse the stronger, deterministic strain in Calvinist thought, but that was rectified by the great *Westminster Confession* of 1646, which did exemplify those convictions in a number of ways. All these things demonstrated that theological determinism is well-grounded in the Reformed tradition.

Then we moved to criticisms of Calvinist position. Between them, Wesley and Channing made the following arguments. (i) That Calvinism removes moral responsibility, because it implies that man cannot do otherwise than God has decreed. This is the argument from ability. (ii) That Calvinism, with its doctrine of unconditional reprobation, is incompatible with God's loving nature. This is the argument from the love of God. (iii) That Calvinism, again on account of unconditional reprobation, is inconsistent with the fatherhood of God. This is the argument from divine paternity. (iv) Lastly, we also saw concerns about the justice of Hell. The chapters that follow are devoted to responding to these objections. Chapter 4 will deal with the ability argument. Chapter 5 will deal with an argument called the manipulation argument, not mentioned here because it is a recent invention, and a problem for all compatibilists. Chapter 6 addresses the question of the justice of Hell. Chapter 7 provides a theodicy of reprobation—explaining what it is that motivates God to reprobate, thereby vindicating his goodness—and Chapter 8 discusses the love and fatherhood of God.

# **PART 2—THEOLOGICAL DETERMINISM AND FREE WILL**

## **Chapter 4—The Ability Argument**

We saw that many of Calvinism’s critics focused their attention on ability claims, and argued that deterministic decrees cancel out moral responsibility because they cancel out the ability to do otherwise. The idea that the ability to do otherwise is necessary for moral responsibility is a popular one—it is called the ‘Principle of Alternate Possibilities’, and is much-discussed in the philosophical literature of the past 50 years. I scrutinise the principle in this chapter. I shall argue that the truth of this principle can be accommodated by theological determinists—I propose that we can use a Kratzer-style semantics of ‘can’ to model ‘could have done otherwise’ statements in such a way that the truth of such expressions is evidently consistent with determinism, and in that way show that Calvinism’s critics were too hasty in assuming that Calvinism ruled moral responsibility by ruling out the truth of such expressions.

### **4.1 Introduction**

As I say, the Principle of Alternate Possibilities (PAP) has been an important principle in the debate concerning free will and determinism. Here is a statement of it drawn from Widerker and McKenna (2006: 2).

PAP: A person is morally responsible for what he has done only if he could have done otherwise.

The principle states a necessary condition for acting in a morally responsible fashion—no alternative possibility, no moral responsibility. But why believe PAP? I see it as a presupposition of a certain conversational practice of ours. Often, when an individual is accused of some sort of misdemeanor and he desires to be exonerated from the accusation, he will protest using language of the following sort: ‘I couldn’t help it!’ or ‘But there was nothing I could have done!’ or ‘I wasn’t able to do anything about it!’. Moreover, it isn’t common to hear a retort of the sort, ‘It doesn’t

matter that you couldn't help it—you are still guilty'. A response like that merely baffles. The far more common way to convince the man of his guilt is to say, 'No, you could have helped it. You could have done such-and-such.' We can see PAP, therefore, as a codification of the principle that governs this practice. We therefore appear to accept that disclamations of the ability to do otherwise (if true) are sufficient for exculpation. But if the absence of this ability is sufficient for the absence of moral responsibility, then the presence of the ability must be necessary for moral responsibility. Hence, PAP.

PAP is an important premise in what I call the Naive Ability Argument for Incompatibilism. This, I take it, is the argument that Wesley et al. have in mind. The argument is as follows:

(1) If someone is morally responsible for performing an action at  $t$ , then they could have done otherwise at  $t$ .<sup>21</sup>

(2) If determinism is true, then, for any time  $t$ , no-one could have done otherwise than they did at  $t$ .<sup>22</sup>

Ergo,

(3) If determinism is true, then no-one is morally responsible for performing any action.

(1) is an expression of PAP, and (2) is the intuitive suggestion that if a description of the past up till  $t$  in conjunction with the laws of nature entails that you perform some action  $A$  at  $t$ , then you could not, at  $t$ , have done otherwise than  $A$ .

Compatibilists are split over how to respond to this argument. Historically, compatibilists denied premise (2); until the 1970s compatibilists were keen to construe abilities in such a way that it was clear that, if their account were correct, then determinism didn't remove the ability to do otherwise. Such compatibilists are now called 'classical compatibilists' (see Berofsky 2006). They typically defended conditional analyses of ability, saying something like the following:

(CAB) An agent  $S$  has the ability to do otherwise at  $t$  iff were  $S$  to try/sufficiently desire/intend to do otherwise at  $t$ , they would do otherwise.

---

21 Advocates of PAP admit that you may be morally responsible in cases of *derivative* responsibility for an action you performed at  $t$  even though you couldn't have done otherwise at  $t$ . Recall that you are usually still responsible for the things you can't help doing when drunk because you usually could have refrained from getting drunk in the first place. It is *non-derivative* moral responsibility is the object of concern here.

22 The first instant of time, if there is such, may be considered an exception to this. But since human beings and their acts came to be after the first instant had been and gone, this point is immaterial. Make  $t$  a variable ranging over all instants later than the first instant, if you wish. See Bailey 2012 for discussion.

How they construed the content of the antecedent varied from one such compatibilist to another, but it is clear that the truth of determinism is not itself any obstacle to our possessing abilities understood in this fashion. The truth of determinism entails that they *won't* try/sufficiently desire/intend to do otherwise; nevertheless, *were* this to happen, they would do otherwise.

But compatibilists fell away from this view. Beginning in the 1970s, a new breed of compatibilist appeared: the semi-compatibilist. With John Martin Fischer as their leader, they grew in numbers and influence until they became the dominant strain in compatibilist thought. The semi-compatibilist responds to the argument by denying (1)—he judges it a mistake to suppose that the ability to do otherwise is necessary for moral responsibility.

What accounts for this changing of the guard? What gave the semi-compatibilist the *chutzpah* necessary to deny such an intuitive principle? Several things. On the one hand, there were Austin (1961) and Lehrer's (1968) attacks on CAB. But the chief catalyst of change came in 1969 with Frankfurt's influential paper 'Alternate Possibilities and Moral Responsibility'. In that paper Frankfurt offered several counterexamples to PAP; a developed version of the most celebrated of them is as follows.

Black is a neuroscientist of considerable expertise who hates Smith. Indeed, Black hates Smith so much that he desires Smith dead. Now Black also knows that Jones hates Smith. One day Black finds out, to his delight, that Jones has formed a plan to murder Smith. But Jones is a temperamental fellow, and Black is worried that Jones might change his mind or that his nerve might fail him. So Black implants a device in Jones's brain without Jones's knowledge. This device monitors Jones's brain activity, and as soon as there is any indication that Jones is not going to follow through on his plan, then Black will use the device to cause Jones to kill Smith. As it happens, Jones shows no sign of renegeing on his plan, and he murders Smith.

Surely it is true, says Frankfurt, that (i) Jones is morally responsible for killing Smith and that (ii) Jones was not, thanks to Black's device, able to do otherwise. So, PAP, says Frankfurt, is false. Semi-compatibilists agree.

After word of these things spread, classical compatibilism fell out of fashion and the semi-compatibilist approach became more popular. Frankfurt changed everything, or at least the face of compatibilism. But I think that was a mistake. I shall defend in this chapter a modest classical compatibilism that assumes that it is not, strictly speaking, *abilities* that PAP is concerned with, but a certain species of *possibility*. My account builds on the Lewis-Kratzer understanding of 'can', and

it permits us another way of seeing that the truth of determinism is compatible with the possibility of doing otherwise.

## 4.2 Analysis of modals

David Lewis, in a famous 1976 paper on the paradoxes of time travel, considers whether we ought to say, of a man who has travelled back in time and who is appropriately situated to kill his grandfather, that he *can* kill his grandfather. On the one hand, we can suppose he satisfies all the ordinary criteria for being able to do it (well-armed, fit, has an excellent opportunity), and therefore that he can; but on the other hand, it appears sound reasoning that he can't, for we can infer from the man's very existence that his grandfather survives and reproduces. Lewis proposes the following resolution: 'To say that something can happen means that its happening is compossible with certain facts. Which facts? That is determined [...] by context. [...] What I can do, relative to one set of facts, I cannot do, relative to another, more inclusive, set.' (1976: 150). So, relative to the fellow's intrinsic properties and opportunity, he can kill his grandfather, but relative to a broader set of facts that includes his grandfather's future existence, he cannot. Disappointingly, beyond the resolution of this paradox, Lewis doesn't go on to deploy this framework as part of a general theory of ability ascriptions.

But Kratzer does. Her approach, like Lewis's, sees context as supplying a set of facts relative to which the 'can' claim is assessed. In her influential 1977 and 1981 papers (reprinted and revised in her 2012a collection), she offers a premise semantics for the modals 'must' and 'can', according to which 'must' functions to express the logical consequence of a proposition from a set of premises and 'can' functions to express the compatibility of a proposition with a set of premises.

By way of illustration, she asks us to consider the following expressions:

- (A) All Maori children *must* learn the names of their ancestors.
- (B) The ancestors of the Maoris *must* have arrived from Tahiti.
- (C) When Kahukura-nui died, the people of Kahungunu said: Rakaipaka *must* be our chief.

The three expressions all contain a 'must' but the 'musts' appear to belong to different modalities. The 'must' of (A) appears deontic and so concerns something like duty; the 'must' of (B) looks like it is epistemic; and the 'must' of (C) looks like it concerns something like prudence: what would be best for the people of Kahungunu.

So, given the apparently different foci of these ‘musts’, what is their common element? Kratzer thinks that sentences (A)–(C) are incomplete. To make them complete we add something like the following:

(A) *Given the duties incumbent on Maori children*, they must learn the names of their ancestors.

(B) *Given what we know*, the ancestors of the Maoris must have arrived from Tahiti.

(C) When Kahukura-nui died, the people of Kahungunu said: *In view of what is best for the tribe*, Rakaipaka must be our chief.

The italicised portions indicate the clauses that were suppressed in the original expressions, what Kratzer calls the conversational background. It is these hidden portions that function to pick out the set of premises that the central proposition asserted is assessed relative to. Following Kratzer’s (2012b) terminology, let’s call this picked out set of premises the ‘modal restriction’. So we can distinguish three elements in the resulting picture: the modal force—a ‘must’ or a ‘can’; the modal scope—the central proposition explicitly asserted; and the modal restriction—a contextually supplied set of premises relative to which the modal scope is assessed.

If we say all that, then we can give an account of the meaning of ‘must’ and ‘can’ that isn’t a long and tiresome disjunctive one that makes reference to every different sort of modality. Instead we can just say this:

‘Must’ expresses that the

<modal restriction>            logically implies the            <modal scope>

and ‘can’, being the dual of ‘must’, expresses that the

<modal restriction>            is logically compatible with the            <modal scope>

It isn’t hard to see how we get the modal scope—it is the central proposition asserted—but what of the modal restriction? How do we arrive at that? The answer lies in the conversational backgrounds. In Kratzer’s view these are contextually provided functions that assign to every possible world a set of propositions (the premises). By way of example, consider (B’) above. The function that context provides is an epistemic one, concerned with what a certain group of people know. So for each

possible world the function assigns a set of propositions (or premises) that include all and only that which is common knowledge at that time for that community in that world.

So with  $f$  as the contextually supplied function from the set of all possible worlds, and  $w$  and  $p$  being variables ranging over worlds and propositions respectively, we can put Kratzer's central idea in symbols (where the ' $\Rightarrow$ ' denotes strict implication).

'Must' expresses that

$$f(w) \Rightarrow p$$

and 'can' expresses that

$$\sim(f(w) \Rightarrow \sim p).$$

This much more or less captures Kratzer's early 1977 work on the topic.<sup>23</sup> But her later 1981 work gave a more complicated picture. She later suggested that the conversational background is really composed of two functions from worlds to sets of premises, namely a *modal base*  $f$  and an *ordering source*  $g$ . The modal base functions, as above, to get us a set of premises, and the ordering source does the same. But the set of propositions derived by the ordering source is used to induce an order on the worlds that the modal base is true at. Then 'must  $p$ ' holds iff  $p$  is true in all worlds closest to the ideal specified by the ordering source, and 'can  $p$ ' holds iff  $p$  is true in at least one of the worlds closest to the ordering source.

I do not wish to go into detail about this later account and the reasons for introducing this more complicated apparatus. I will, for ease of understanding and accessibility, work with the framework given by Kratzer's earlier account. However, for those familiar with Kratzer's later work, I will make it plain in a later footnote how my theory should be stated relative to her more developed account.

### 4.3 Analysis of 'could have done otherwise'

We now need to apply all this to PAP. PAP is the claim that the truth of the expression 'S could have done otherwise' is necessary for S to be morally responsible for whatever act S performed. But

---

<sup>23</sup> I have left out how she handles inconsistent modal restrictions, but that is immaterial for my purposes because, on the account I will give, it will not be possible for the modal restriction to be inconsistent.

‘could have done otherwise’ is an ambiguous expression in at least two ways. Firstly, ‘could have’ is ambiguous between a subjunctive and an indicative reading.<sup>24</sup> On the indicative reading it occurs as the past form of ‘can’, where ‘can’ is understood to mean ‘be able to’, such that ‘S could have done otherwise’ means ‘S was able to do otherwise’.<sup>25</sup> On the subjunctive reading, ‘could have’ is the past form of the present tense subjunctive ‘could’. This subjunctive interpretation doesn’t have abilities in view, and instead holds that it is various circumstances that might have obtained that are in view, various possibilities that could have come to pass. On such a reading, the ‘could have’ can be viewed as equivalent to the non-epistemic ‘might have’, such that ‘S could have  $\phi$ -ed’ means ‘S might have  $\phi$ -ed’, rather than ‘S was able to  $\phi$ ’.

So, the indicative ‘could have’ deals with ability, and the subjunctive ‘could have’ talks about possibility. But once we are aware of this, the following idea might occur to one: rather than assuming that PAP is concerned with the *ability* to do otherwise, we might instead suppose that it is concerned with the *possibility* of doing otherwise. We could assume that it is only the truth of the subjunctive ‘could have’ that is necessary for moral responsibility, while supposing that the truth of the indicative ‘could have’ is not, strictly speaking, necessary. In this chapter, I shall make this assumption. I say: grant me this modest assumption, and I can give you a classical compatibilism that evades the standard objections.

A great advantage of making this move is that one avoids entirely the debate about satisfactory analyses of ability ascriptions. Abilities aren’t in view at all, only a certain sort of possibility.

But is it a plausible move to make? Someone might make the following objection: ‘But we don’t talk only about whether or not someone could have done otherwise, we also talk about, in present cases, whether some individual can do otherwise, and ‘can’ is in the indicative.’ It is true that ‘can’ takes the indicative, but it would be a mistake to think that, because of this, it must always be abilities rather than possibilities in view. A distinction is often drawn between the “‘can” of ability’ and the “‘can” of possibility’ (see Vetter 2015: 76). That there is a clear possibilist use of ‘can’ can be seen from the following examples: ‘You got burgled? Well, these things can happen.’ ‘Learning a language can be a difficult affair.’ ‘Propositions can be true; contradictions cannot be true.’ They all resist a parsing into ability language. ‘Burglaries are able to happen.’ ‘Learning a language is able to be a difficult affair.’ ‘Propositions are able to be true, but contradictions aren’t able.’ All such expressions, even if one might not be prepared to say that they are strictly incorrect, are nevertheless decidedly awkward and unnatural. This is because we are trying to get ability-talk

---

24 A distinction van Inwagen ably deploys in his 1984 response to Dennett 1984.

25 Or perhaps, more carefully, ‘S was both able and had an opportunity to do otherwise’, if we think that the ‘have’ modifier forces our attention to a particular occasion. ‘I could run a marathon’, where ‘could’ is the past form of the abilitative ‘can’, talks about a general ability possessed in the past; ‘I could have run a marathon’ suggests that one was able and also in a good position to run a marathon on a particular occasion. I thank Simon Kittle for this point.

to do a job better done by possibility-talk. One way of detecting the ‘can’ of possibility is to see whether the sentence carries an identical meaning if you replace the ‘can’ with ‘may’ (just so long as it is clear that the ‘may’ is not expressive of permission or the deontic modality). This works for the above examples. ‘Burglaries may happen.’ ‘Learning a language may be a difficult affair.’ ‘Propositions may be true, but contradictions may not.’ This is because ‘may’ can also be used to express possibility.

So, I can accommodate currently occurring moral-responsibility contexts by supposing that the ‘can’ in the ‘can do otherwise’ that occurs on such occasions is the ‘can’ of possibility. It might be suspected that I press for possibility rather than ability because I think that Kratzer’s apparatus, for all its success, cannot satisfactorily handle abilities.<sup>26</sup> I agree that Kratzer’s account appears in its best light when it is seen as giving a model of the ‘can’ of possibility, rather than the ‘can’ of ability. Nevertheless, whether or not the Kratzer semantics can give an adequate analysis of ability ascriptions, discussion of that remains for another occasion. Defending that supposition would involve, among other things, responding at length to the formal objections made against the idea by Anthony Kenny (1975: 136–7).

It might also be objected, however, that we also use ‘able’-talk when we talk about whether or not someone ‘can do’ or ‘could have done’ otherwise. We also ask ‘are they *able* to do otherwise?’ and ‘were they *able* to do otherwise?’. We do indeed use such expressions, but, on the view I am proposing here, this is a mistake, though a very understandable one. Given the very close connection between the ‘can’ of ability and the ‘can’ of possibility, and the subjunctive ‘could have’ and the indicative ‘could have’, we can’t expect the ordinary-language user to pay careful attention to such subtleties, especially in a context that concerns agents. Indeed, we would positively expect him to slide, in his benighted ignorance, from possibility-talk to ability-talk and back again, thus confounding what should be kept separate.

One last objection that might be made at this juncture is this: if I am eschewing abilities as, strictly speaking, unnecessary for moral responsibility, then it might be wondered if the account I am giving is really classically compatibilist at all. Weren’t the classical compatibilists of old all concerned to defend the relevance of ability ascriptions to moral responsibility? Well, I agree that they were, but it is also nevertheless clear that I am a classical compatibilist of some sort. I accept PAP, and therefore accept that the truth of a ‘could have done otherwise’ expression is necessary for moral responsibility. Moreover, I don’t say anything like, ‘People naively supposed that alternative possibilities were necessary for moral responsibility. But that was a mistake. What we are *really*

---

<sup>26</sup> Maier’s pessimistic judgement that ‘the Kratzer semantics alone does not suffice to settle questions about the agentive modalities.’ (2013: 115) appears to be a popular one.

interested in when we ask whether or not someone could have done otherwise is the following feature ...'. That is the sort of thing semi-compatibilists say. I am not to be counted amongst their number. I accept the requirement for alternative possibilities. That said, I recognise that the account I am giving is not as fully fledged or as substantial as other classical compatibilisms—but that is why I call it 'modest'.

Finally, let's return to the second way in which 'could have done otherwise' is ambiguous: we don't know what doing otherwise involves. I shall take it to mean intending otherwise. I think examples like Fischer and Ravizza's 'Sharks' case show this (1998: 125). John is walking by a pool. He sees a child drowning. He is not inclined to get himself wet, so he continues walking. We think John is a wicked man, and we are inclined to blame him for not saving the child. However, unbeknownst to John, there were sharks in the water, and, were John to have attempted to rescue the child, he would have been set upon and eaten. So, he couldn't have saved the child in any case. Does this mean we cannot blame him? After all, he couldn't have done otherwise than fail to save the child. But I think that only shows that he must be guilty on account of the presence of a different alternative possibility, namely, that he could have *decided* to try and save the child yet did not. But such decidings and choosings are, as I supposed at the start of this thesis, simply the forming of intentions. So, we can hold John guilty for not *intending* otherwise than he did.

#### **4.4 Applying the Kratzer semantics to 'could have done otherwise'**

So, if we are understanding 'could' in the way proposed by Kratzer's early model, then we need to ask ourselves: what is the modal restriction? It is clear what the modal scope is: it is 'S intends otherwise at  $t$ '. But if 'could' expresses the compossibility of a set of premises with the modal scope, then we need to know what these premises are. What sort of conversational background does our 'could have done otherwise'-talk presuppose? If we use ' $f$ ' as a variable ranging over functions that perform the role of the modal restriction, then we can use ' $f_m$ ' as a variable ranging over a certain subset of those functions, which subset handles the sort of moral-responsibility contexts we are concerned with.  $f_m$ , I shall suppose, will vary according to which particular occasion of an agent's acting is being considered. For every possible world in which the agent in question acts on that one particular occasion in question,  $f_m$  will assign to that world a non-empty set of propositions.  $f_m$  is therefore a function from worlds to sets of premises; but what are these premises to which it takes us?

If we are to cast the incompatibilist in a Kratzerian mould, then he might say this:  $f_m$  takes from every possible world two things in forming the premise set; it takes (i) the laws of nature that hold at that world (N) and (ii) a complete description of the history of the universe of that world that includes all and only such moments that are prior to the agent performing his act at  $t$  (D).<sup>27</sup> And, indeed, if it is true that

(4)  $\{N, D\} \Rightarrow \sim(S \text{ intends otherwise at } t)$ ,

it therefore follows, if the incompatibilist is right, that S could not have intended otherwise relative to this set of premises.

But we don't need to take the incompatibilist's word for it that this is how to characterise the function in question. It is my contention that we should take  $f_m$  to be a function from worlds to propositions describing psychological setups such that the relevant possibility of doing otherwise consists in something like *psychological compossibility*—the idea that S's intending otherwise at  $t$  must be compossible with S's psychological setup prior to  $t$  in order for the morally relevant alternative possibility to obtain. What do I mean to include in these psychological setups? Let me explain.

I introduce my account by way of the following distinctions made by Robert Kane concerning the ambiguity of the term 'will'. Kane says we ought to distinguish the following three senses of the word 'will':

- (i) what I *want, desire, or prefer* to do
- (ii) what I *choose, decide, or intend* to do
- (iii) what I *try, endeavor or make an effort* to do (Kane 1998: 26)<sup>28</sup>

He calls (i) the desiderative or appetitive will, (ii) the rational will, and (iii) the striving will (1998: 27). The desiderative will is easy to understand: it is your desires, your wants. The rational will concerns the decisions you make, typically (at least in part) on the basis of your desires; if you like to drink coffee, you may decide to purchase some. The distinction between the rational and the striving will is more subtle, but it is recognised once we realise that deciding to buy some coffee isn't sufficient for striving for some coffee. I may decide to buy some coffee while I am at work, but

---

27 Though an incompatibilist such as van Inwagen (1984) would, if I read him correctly, take the question of the compossibility of N and D with doing otherwise as a *consequence* of his account of ability, rather than as an analysis of it.

28 To Kane's list we might also add commands as a type of will. To flout the commands of a monarch is to flout his will, even though the monarch might not desire or intend, for some reason(s), that his commands be fulfilled.

because I know that I won't have the opportunity to buy some coffee until I leave work, I only try or endeavour to buy some coffee after I leave the office.<sup>29</sup>

Kane also says, 'If there is indeterminacy in free will, on my view, it must come somewhere between the input and the output—between desiderative and rational will.' (1998: 27). And this, I think, is right. I think all parties to the debate, whether compatibilist, libertarian or whatever, must acknowledge this threefold distinction of will. Moreover, we all also realise that the locus of the free-will debate is found here between the desiderative will on the one hand, and the rational and striving will on the other. As Kane points out, I think we would all be happy with a purely deterministic relation from the rational to the striving will, but we are not all happy with a deterministic relation from the desiderative will to the rational will, for libertarians would take strong exception.

But note that our judgements of praise and approbation are sensitive too to this distinction between the desiderative will and the rational will. If a man is genetically disposed to a short temper, then we don't blame him for that. We say something like, 'He can't help the desires he was born with [the desiderative will], but we can hold him accountable for what he does with those desires [the rational will] (so long as they don't compel him)'. The distinction here concerns what an agent is given, what he finds himself with, and what he does with what he has been given. The desiderative will concerns the former, and the rational will the latter. This distinction is one we all intuitively acknowledge in our assessments of moral responsibility.

But once armed with this distinction I think we have all we need to state a plausible classical compatibilism. We say that  $f_m$ —the modal restriction found in PAP—is a function that moves from a world to a set of premises that offers a complete description of the relevant agent's desiderative will prior to their choosing. The central suggestion of this chapter can therefore be put like this: to claim that 'S could have done otherwise' in the sense required for moral responsibility is just to claim that S's intending otherwise was compatible with their desiderative will. More formally, that

(5)  $\sim(\text{S's desiderative will} \Rightarrow \sim(\text{S intends otherwise}))$ .

And to say that 'S couldn't have done otherwise' means

(6)  $\text{S's desiderative will} \Rightarrow \sim(\text{S intends otherwise})$ .

---

<sup>29</sup> It may be that the distinction between the rational and the striving will doesn't amount to much. It may be only the distinction between deciding now to do something now and deciding now to do something later.

I think we can see the force of this idea when we consider scenarios where we would be inclined to describe the desiderative will as entailing the rational. Consider a case of a man with a strong addiction to heroin; in fact, his addiction is so strong that it is literally irresistible. We don't wish to hold a man responsible for what he does under irresistible compulsion, and my account explains why: in such a case it looks as if what the agent chooses follows inevitably from his desiderative will, his addictive desires. Intuitively speaking, these make choosing otherwise impossible relative to the desiderative will and so remove his moral responsibility.

So far so good, but in fact (5) and (6) are too weak. A description of the desiderative will by itself doesn't really entail very much. We want it to be the case that the addict's decision is entailed by his desiderative will. But it won't be, for without any beliefs about how to fulfil one's desires, even an overwhelming desire can find no outlet in action. So we should include a description of the agent's doxastic states alongside the description of their desiderative will. For the sake of completeness, we might also want to include the agent's experiential states as well, in case we think they might have some bearing. I will suppose that the doxastic, desiderative and experiential states of the agent constitute the agent's *psychological setup*. I intend the psychological setup to include all those aspects of the agent's makeup that are commonly supposed to be the 'springs of action'—beliefs and desires, even if they are not enough to provide an analysis of decision, are often considered adequate to account for all those features explanatorily relevant to the agent's decision-making in ordinary cases.<sup>30</sup> So, we should therefore replace 'S's desiderative will' in (5) and (6) with 'S's psychological setup'.

Is the psychological setup enough to get us the required entailment from addictive desires to acting upon such desires? It is not, because it may be that the addict ceases to exist before any choice issues from his mental faculties. So, we need to include S's continued existence (E) alongside a description of S's psychological setup. But including S's continued existence in (5) and (6) isn't enough either. For suppose that an addict is going to act on an irresistible desire. We want it to be impossible for him to do otherwise. However, again, it won't be, because it may be that an external force (an angel swooping down from heaven, say) interferes and causes the addict to intend otherwise than that which the desiderative will would have determined. So, we must also include the claim that the agent is not interfered with in such a way. All I intend to include in this non-interference claim (NI) is the idea that no force external to the agent and the agent's psychological setup is directly causally responsible for the agent's producing his decision. The only admissible

---

<sup>30</sup> I am not particularly concerned to limit the psychological setup of the agent to beliefs, desires and experiences. If it should emerge that it would be better for my account to include yet more aspects of the mind as part of the psychological setup, then I am happy for the psychological setup to be expanded as required—with one proviso: that it does not include what I introduce below as the 'individual nature of the agent'.

candidates are features intrinsic to the agent or his psychological setup. But continued existence (E) and non-interference (NI) aren't enough either; we must also include the laws of psychology (P), for without these there is nothing to rule out the possibility of an overwhelming desire to  $\phi$ 's giving rise to  $\psi$ -ing rather than  $\phi$ -ing.

The resulting analysis is therefore more complicated. Suppose S made a decision at  $t$ . Strictly speaking,  $f_m$  must generate all the above premises such that a more complete understanding of 'S could have done otherwise at  $t$ ' in the sense required for moral responsibility is this:

$$(7) \sim(\{S\text{'s psychological setup, E, NI, P}\} \Rightarrow \sim(S \text{ intends otherwise at } t)).$$

Spelled out:

An agent S could have done otherwise at  $t$  (in the sense required for moral responsibility) iff a description of S's desiderative will, and S's doxastic and experiential states, over an interval<sup>31</sup> immediately prior to  $t$  in conjunction with the claims that (i) S continues to exist until  $t$  and (ii) S's decision-making at  $t$  was not directly caused by any force external to the agent or his psychological setup, and with (iii) the actual laws of psychology, do not entail that S choose as he did at  $t$ .<sup>32</sup>

Is this the final analysis? I think we must make one more adjustment. (7) would give, I think, the wrong result in cases where the decision of the agent is the result of direct causal interference by an external force. If an angel swoops down from heaven and causes me to decide, in a manner that bypasses the psychological setup, to sing 'God Save the Queen', then wouldn't I, if people complained about the noise and I were aware of what the angel had done, complain that I could not help it? But (7) would get the result that I could have helped it, because my not choosing to sing would be entirely consistent with my psychological setup, etc. just prior to the angel swooping down and causing me to decide to sing.

As a result, I think we must also suppose NI to be actually true, and the decision brought about in the proper manner for the 'could have done otherwise' claim to be true. So, in addition to the content generated by  $f_m$ , we will have to add the non-interference claim alongside it. The final analysis of 'S could have done otherwise at  $t$ ' is therefore:

---

31 It does not matter which interval, for, even if the interval is very long, what S's psychological setup was like in the distant past won't be relevant to what it is possible for S to intend at  $t$ —what will be relevant is S's psychological setup as  $t$  approaches.

32 For those desirous to know how my account would be incorporated into Kratzer's later framework, it is achieved as follows: let the propositions detailing S's instantiation of his psychological setup be the modal base. E, NI and P are to be considered as the ordering source.

(8)  $NI \ \& \ \sim(\{S\text{'s psychological setup, E, NI, P}\} \Rightarrow \sim(S \text{ intends otherwise at } t))$ .

‘S couldn’t have done otherwise at t’ is just the negation of (9):

(9)  $\sim(NI \ \& \ \sim(\{S\text{'s psychological setup, E, NI, P}\} \Rightarrow \sim(S \text{ intends otherwise at } t)))$ .

It might be thought that narrowing things down to whether or not an agent can intend otherwise given his psychological setup is too restrictive a model to handle the great variety of moral-responsibility contexts.<sup>33</sup> Suppose a mugger holds a gun to my head and suggests it would be prudent for me to relinquish possession of my immediate finances. I agree and hand my wallet over. If, after the fact, someone were to challenge the propriety of my doing so, I might well respond that ‘I couldn’t have done otherwise’. Such an expression would be strictly false on my account: after all, I could have formed the intention to fight the mugger. Or suppose I am tied up and a child drowns in front of me as I watch helplessly. I would again complain that I couldn’t have done otherwise. But this will likewise be false on my account: I could have intended differently.

One might think, therefore, that an account of the modal restriction that varies more widely in generated content given context would be desirable. In the mugging case, one might say, the modal restriction wouldn’t generate a description of my desires, but instead a set of rules about what are reasonable courses of action given the value of one’s life. In that case, what I mean when I say ‘I couldn’t have done otherwise’ is that my resisting the mugger was not possible relative to those rules. In the case of the drowning child, it might be suggested that the modal restriction would generate a list of all possible bodily exertions in that situation, such that the ‘I couldn’t help it’ claim expresses the conviction that relative to my position and the laws of nature, there was no possible bodily exertion that would have resulting in the breaking of my bonds.

Such an account might, at the end of the day, be a better way to go. There would be no type of modal restriction distinctive to moral-responsibility contexts in that case. Nevertheless, my more uniform account can still handle these sorts of examples. In the mugging case, what I really mean when I say that ‘I couldn’t have done otherwise’ is that I couldn’t have *reasonably* done otherwise. In other words, yes, I could have done otherwise, but it wouldn’t have been reasonable to do so (because I would have been shot). In the drowning child case, when I say that ‘I couldn’t have done otherwise’ I am again speaking loosely. Indeed, strictly speaking, I could have done otherwise, but

---

33 I thank Pablo Rychter for this suggestion.

what I mean is that, of all the things I could have intended, none of them would have had the desired result (the rescuing of the child).

Whatever the advantages of variant accounts, it is the analysis given in (8) and (9) that I am defending on this particular occasion. It seems to get us what the compatibilist wants. It appears sufficient to ensure that the addict cannot do otherwise, yet doesn't require anything like the falsity of determinism. Furthermore, it avoids all the old objections to CAB. I have not analysed the 'could have done otherwise' claim as a subjunctive conditional, and therefore the objections to the old classical compatibilism aimed to exploit this feature cannot be deployed against my proposal. I have made the suggestion that the alternative possibilities requirement on moral responsibility is concerned with *compatibility* or *compossibility*, rather than the truth of various 'would' conditionals: broadly speaking, just so long as one's desiderative will is compatible with one's deciding to do otherwise, then one 'could have done otherwise' in the sense required for moral responsibility. Such, I suggest, is the compossibility that PAP is fundamentally concerned with.

#### **4.5 Frankfurt and other considerations**

Why might compatibilists have overlooked this suggestion? I think there are two reasons. Firstly, compatibilism has often been understood, both by its proponents and its opponents, as precisely the claim that the rational will is fixed by the desiderative will! The traditional 'strongest desire wins' compatibilism of Hobbes (1651) and Edwards (1977 [1754]), for instance, is a compatibilism that insists that our decisions are determined by what the understanding recognises as the strongest desire. I don't think that such compatibilisms have an adequate way of handling overwhelming desires—the addict, if literally overwhelmed by his desire for a drug, surely has his responsibility diminished, if not removed—so I don't consider it a problem that my way of securing PAP is inconsistent with their proposals.

Secondly, it might be thought that the phrasing I reach for is too close to libertarian phraseology for comfort. Locating the freedom of the agent in what the agent does with his desires and so forth looks like exactly the sort of thing libertarians have demanded throughout the years. Well, I would concede that the theory I give here is a concession of sorts to the libertarian: he was quite right to insist on a measure of independence for the rational will from the desiderative. It might then be wondered how I can concede this consistently with the truth of determinism. For if the desiderative will, or the psychological setup more broadly, doesn't determine the rational will, then what does? But what lies between the psychological setup and the rational will? It is the agent. It is he who

decides what is going to happen given the content of the desiderative will. Accordingly, I move that it is the individual nature of the agent that determines the rational will. I don't mean by 'nature' the essential properties of the agent—they might not be essential to either being an agent or being that particular agent—I merely intend to describe categorical properties of the agent distinct from the psychological setup that I think explain why one decision rather than another is issued in typical cases of decision-making.

This helps make it clear how my theory differs from the libertarian's (or Arminian's); for let 'A' denote S's individual agential nature; then I am happy to agree that

(10) {S's psychological setup, E, I, P, A}  $\Rightarrow$   $\sim$ (S intends otherwise at *t*)

can be true consistently with the truth of 'S could have done otherwise' (in the moral-responsibility sense). But the libertarian denies this. For the libertarian insists that, broadly speaking, there must be no entailment from the psychological setup plus *anything* (or at least any hard fact<sup>34</sup>) to the rational will for the relevant possibility of doing otherwise to obtain. I contend for the weaker position that, speaking loosely, there must merely be no entailment from the psychological setup to the rational will. Nevertheless, still speaking loosely, I do believe that there is an entailment from the psychological setup *plus* the nature of the agent to the rational will.

Finally, let's deal with Frankfurt's celebrated counterexample to PAP. Frankfurt was trying to give us an example where an agent was morally responsible for what he did, but couldn't have done otherwise. I think it should be clear, relative to my account of 'could have done otherwise', that Frankfurt's scenario is one where Jones *could* have done otherwise. (8) is a conjunction, and so if Frankfurt's counterexample is to succeed, then at least one of these conjuncts will have to be false of Jones. Is the non-interference (NI) claim false? It is not. It is an important part of Frankfurt-style counterexamples that there is no actual interference by Black or his device at all. Such things turn out to be entirely unnecessary in securing Jones's decision. Is it the second conjunct that is false? Is there any entailment from Jones's psychological setup, in conjunction with his continued existence, his not being interfered with, and the laws of psychology to his not doing otherwise than deciding to go ahead and kill Smith? We have no reason to think so. We would have, if Jones acted from an addiction or compulsion. But again, Jones's decision is supposed to be (aside from its grim objective) a regular and ordinary piece of decision-making. Jones is not supposed to be driven along by an overwhelming inferno of desire. We do not, therefore, have any reason to think that his decision to go ahead and kill Smith is entailed in the relevant way, quite the contrary.

---

34 See Plantinga 1986 for an explanation of the distinction between hard and soft facts.

Nor does Black's brooding presence over Jones's activity do anything to affect this lack of entailment. The only way in which we could get the conclusion that Jones couldn't do otherwise (in the relevant sense) on account of Black is if the set of premises  $f_m$  points us to includes a description of Black, his behaviour, his device, and his motives. But my account is of a fundamentally abstractive character: I am interested in whether or not the agent's psychological setup, *considered in the abstract*, is consistent with a different intention being formed. So,  $f_m$  will never stray so far outside the agent as to include someone else and their schemes in the set of premises it points us to, and for this reason Black will always be irrelevant, and all such Frankfurt-style counterexamples will fail to counter.

## 4.6 Conclusion

Thus, we see that there is no need for the Calvinist to insist, as Luther appeared to do, that because it is impossible for the reprobate to turn to Christ, and because the reprobate is nevertheless blamed for his refusal to turn, our moral-responsibility discourse is radically mistaken. It is indeed impossible *relative to God's determining decree* that the reprobate should turn, but it is *not* impossible relative to the reprobate's psychological setup, and it is this fact that permits the Calvinist to retain, without substantive alteration, the natural assumption that blameworthiness requires that one 'could have done otherwise'.

So much for the argument from ability. I turn in the next chapter to what is called the 'manipulation argument'. This argument is a recent development, and therefore not to be found on the lips of the historical opponents of Calvinism. But the argument's great place in recent discussions of compatibilism and incompatibilism make it prudent for it to be addressed.

## Chapter 5—The Manipulation Argument

### 5.1 Introducing the Argument

For the reasons given in the past chapter, I am largely unmoved by the ability-based arguments for incompatibilism. I am more impressed by manipulation arguments. There is, I feel, a bite to manipulation arguments that is absent from arguments from ability. Moreover, as we shall see, the manipulation argument threatens to have extra bite when directed at the theological determinist. There are objections to the argument that the secular determinist can make that his theistic counterpart cannot. But we shall also see that there are unexpected advantages that the theological determinist will have when crafting a response that his secular peer does not have access to.

What is the argument, then? Consider the following scenario. You work a dreary old job in a dreary old office. But one day at your place of work something interesting happens. Your colleague, Ivan, is peculiarly animated this morning. He gestures you over to his desk. You comply. Here shows you a very intricate piece of technology that he is hiding beneath his desk. ‘With this device,’ he whispers, ‘I can tap into people’s brain waves. More than that, I can send signals to people’s brains, causing them to do whatever I program into this device!’ You are incredulous. ‘Don’t believe me, eh?’ He replies. ‘Well, you just watch this.’ He fiddles with the device. ‘I’ve just tapped into the boss’s brain waves. You remember those briefs we handed in late last week? I’ve send instructions to his brain that will cause him to storm out of his room in a few seconds, and rant at us for handing them in late.’ Sure enough, a few seconds later, the boss emerges from his office, red-faced and furious, and proceeds to harangue you and your colleague at length for your tardiness, before returning to the privacy of his studio. You are amazed. ‘Perhaps it is just a coincidence’, you suggest falteringly. ‘Still unpersuaded?’ says Ivan, fiddling once more with the device. ‘Fine. I’ve just send another signal to the boss’s brain. This time I’ve told him to come out and sing us some opera.’ Once again, the incredible happens. Your boss emerges from his enclosure, and the whole office floor is treated to a (surprisingly pleasant) rendition of the anvil chorus from Verdi’s *Il Trovatore*.

Were such a scenario to come to pass, you would no doubt be greatly astonished that the decisions of another human being could be so precisely manipulated. But that is not the reaction we are interested in. We are interested in how your assessment of the boss’s culpability for his actions is affected by the knowledge that Ivan caused the boss to act in that way through his device.

Suppose another member of the office overheard the boss's rant and complains to the boss's boss, and the boss is summoned to respond to accusations that his behaviour crossed professional lines. Our intuitive reaction is that the boss is unfairly accused. 'After all,' we want to say, 'it wasn't his fault. It was Ivan's fault. Ivan *made* the boss do it. It wasn't the boss; it was Ivan. The boss was merely Ivan's programmed tool. You can't be responsible if you are a mere tool, *made* to do something by someone else, can you?'<sup>35</sup>

In short, we intuitively feel that the boss is not responsible for actions performed as the result of manipulation of the sort Ivan employed. What does this have to do with compatibilism and incompatibilism? It is an argument for incompatibilism in the following way. *Why* is it that the boss is not responsible for his manipulated actions? The incompatibilist has a straightforward answer: it is because he was caused to do them. He was caused by Ivan's device. An if an agent is caused to perform an action, then that action is neither free nor morally responsible. But the compatibilist can reach for no such explanation—for he thinks that moral responsibility is compatible with an agent's act being determined. It looks like the compatibilist is committed to thinking that manipulated agents are responsible. But that is very counterintuitive, and therefore a good reason to be an incompatibilist over a compatibilist.

The compatibilist will not, of course, think that just any old determining is compatible with moral responsibility. No modern-day compatibilist thinks that actions caused by compelling desires are fit to blame agents on account of. Fischer and Ravizza (1998), for instance, insist that the decision must arise from an appropriately reasons-responsive mechanism. And so on and so forth. But it is the great advantage of the manipulation argument that its advocates can, it appears, simply suppose that all the compatibilist conditions on moral responsibility are met. Because the compatibilist is committed to thinking there is no problem, in principle, with an agent's free action being the inevitable result of a certain chains of causation (chains of causation that satisfy the compatibilist requirements on moral responsibility), there doesn't seem to be any way of blocking off the possibility that the manipulation might come through those same chains of causation. The advocate of the manipulation argument can simply suppose that the manipulating device or person causes the subject's action in the same way, and not through irresistible desire, or through bypassing the agent's deliberative faculties, or what have you.

---

35 In this way, we can see that both of the central arguments for incompatibilism are based on two distinct types of language we employ to perform an exculpatory function. We employ ability- or possibility-based language like 'But I couldn't help it!' to justify the ability to do otherwise as necessary for moral responsibility; and we use force- or cause-based language like 'But he made me do it!' to justify the incompatibility of manipulation with moral responsibility.

## 5.2 Mele's Zygote Argument

Having thus summarised the manipulation argument, I want to lay out the versions of the argument that are discussed in the current literature. Most of the discussion takes place around the two most prominent versions of the argument: Mele's Zygote argument and Pereboom's Four-Case argument. Here is Mele on the former. Diana is a 'supremely intelligent being' (2006: 184), a goddess.

Diana creates a zygote *Z* in Mary. She combines *Z*'s atoms as she does because she wants a certain event *E* to occur thirty years later. From her knowledge of the state of the universe just prior to her creating *Z* and the laws of nature of her deterministic universe, she deduces that a zygote with precisely *Z*'s constitution located in Mary will develop into an ideally self-controlled agent who, in thirty years, will judge, on the basis of rational deliberation, that it is best to *A* and will *A* on the basis of that judgment, thereby bringing about *E*. If this agent, Ernie, has any unsheddable values at the time, they play no role in motivating his *A*-ing. Thirty years later, Ernie is a mentally healthy, ideally self-controlled person who regularly exercises his powers of self-control and has no relevant compelled or coercively produced attitudes. Furthermore, his beliefs are conducive to informed deliberation about all matters that concern him, and he is a reliable deliberator. So he satisfies a version of my proposed compatibilist sufficient conditions for having freely *A*-ed. (2006: 188)

Mele goes on to offer a 'skeleton form' (2006: 189) of the manipulation argument that would be deployed on the basis of such a case:

1. Because of the way his zygote was produced in his deterministic universe, Ernie is not a free agent and is not morally responsible for anything.
2. Concerning free action and moral responsibility of the beings into whom the zygotes develop, there is no significant difference between the way Ernie's zygote comes to exist and the way any normal human zygote comes to exist in a deterministic universe.
3. So determinism precludes free action and moral responsibility (2006: 189)

One should note that the focus has shifted when we come to the numbered argument. It has moved from Diana's intent to bring about *E* via Ernie's *A*-ing (and, therefore, presumably, calling into question Ernie's responsibility for *A*-ing) to questioning Ernie's responsibility for anything (presumably because we are to understand that, in fact, all of Ernie's actions throughout his life have been planned out by Diana). But either the original scenario or the numbered argument can be adjusted to bring the two into uniformity.

There is some unfortunate phraseology in the first premise of the argument. The preliminary clause ‘Because of the way his zygote was produced in his deterministic universe’ has led to some confusions over how to interpret the argument. Kearns (2012), for instance, has wondered what ‘the way’ refers to. He objects that if ‘the way’ refers to deterministic causation, then the first premise will simply beg the question against the compatibilist. On the other hand, if it refers to the manipulative causation of Diana, then the worry emerges that the generalisation from manipulative causation as responsibility-undermining to determinism *simpliciter* as responsibility-undermining will fail. If it is by virtue of *manipulative* causation that Ernie is not responsible, then the problem is to do with manipulation, the compatibilist might aver, not determinism.

But we can largely sidestep such issues. As Patrick Todd suggests, we can do away with the ‘because’ clause. It isn’t necessary. As he proposes, we can replace the first premise with merely

(1\*) Ernie is not free or morally responsible with respect to performing *A* or bringing about *E*. (2013: 193)

Support for premise (1\*) might come from some detailed sub-argument but in typical presentations of the manipulation argument, that isn’t supposed to be how it goes. One’s natural, intuitive reaction to the Zygote case is that Ernie is not free in his *A*-ing. The support for (1\*) can therefore come straightforwardly from bare intuition. This isn’t quite sufficient to remove the worry that something is question-begging. Only those who do not affirm a certain sort of compatibilism can consistently accept (1\*). But we can remove that worry by altering the premise again:

(1\*\*) Intuitively, Ernie is not free or morally responsible with respect to performing *A* or bringing about *E*.

There is no problem with all parties granting that claim. After all, one can consistently grant that one has an intuition that *p* while denying that *p*. The rest of the argument can then proceed as one would expect:

(2\*) There is no responsibility-relevant difference between Ernie’s case and the case of determined agents acting in a way that satisfies compatibilist conditions for moral responsibility.

(3\*) Therefore, intuitively, determined agents acting in a way that satisfies compatibilist conditions for moral responsibility are not morally responsible.

That is the form of the argument that I shall work with.

### 5.3 Pereboom's Four-Case Argument

Pereboom's argument is more complicated. His strategy is this: to present you with a case where it is hopefully obvious that the individual is not morally responsible, and then to proceed, step by step, to cases which come closer and closer to resembling our own case (assuming our universe is deterministic), in the hope of showing that our case is really no different from the first, and we are not morally responsible agents either. He moves across four cases, and the thought is that the difference between the cases is too slender for one to plausibly say that responsibility fails in some case  $x$  (ranging over 1–3), but obtains in case  $x + 1$ .

Pereboom describes the first case as follows:

Case 1. Professor Plum was created by neuroscientists, who can manipulate him directly through the use of radio-like technology, but he is as much like an ordinary human being as possible, given his history. Suppose these neuroscientists 'locally' manipulate him to undertake the process of reasoning by which his desires are brought about and modified—directly producing his every state from moment to moment. The neuroscientists manipulate him by, among other things, pushing a series of buttons just before he begins to reason about his situation, thereby causing his reasoning process to be rationally egoistic. Plum is not constrained to act in the sense that he does not act because of an irresistible desire—the neuroscientists do not provide him with an irresistible desire—and he does not think and act contrary to character since he is often manipulated to be rationally egoistic. His effective first-order desire to kill Ms. White conforms to his second-order desires. Plum's reasoning processes exemplify the various components of moderate reasons-responsiveness. He is receptive to the relevant pattern of reasons, and his reasoning processes would have resulted in different choices in some situations in which the egoistic reasons were otherwise. At the same time, he is not exclusively rationally egoistic since he will typically regulate his behaviour by moral reasons when the egoistic reasons are relatively weak—weaker than they are in the current situation. (2001: 112–113)

The Plum of case  $x$  will be referred to as 'Plum <sub>$x$</sub> '. Case 1 therefore informs us about Plum<sub>1</sub>. Plum<sub>1</sub> is subject to *ongoing manipulation*. He is constantly monitored and when it looks a decision is going to occur that the neuroscientists do not want to occur, his psychology is interfered with so that the decision the neuroscientists prefer is caused. But compatibilist conditions on morally responsible

decision-making are of course satisfied. Plum<sub>1</sub> ends up murdering Ms. White, all thanks to the careful manipulation of these scientists. Pereboom thinks it is obvious that we should not blame Plum<sub>1</sub> for that. He then presents his next case.

Case 2. Plum is like an ordinary human being, except that he was created by neuroscientists, who, although they cannot control him directly, have programmed him to weigh reasons for action so that he is often but not exclusively rationally egoistic, with the result that in the circumstances in which he now finds himself, he is causally determined to undertake the moderately reasons-responsive process and to possess the set of first- and second-order desires that results in his killing Ms. White. He has the general ability to regulate his behavior by moral reasons, but in these circumstances, the egoistic reasons are very powerful, and accordingly he is causally determined to kill for these reasons. Nevertheless, he does not act because of an irresistible desire. (2001: 113–114)

Plum<sub>2</sub> is different because, although, like Plum<sub>1</sub>, he was created by neuroscientists, Plum<sub>2</sub> is not subject to any ongoing manipulation. Plum<sub>2</sub> is programmed to from the start to develop into an individual who will kill poor Ms. White. He is subject to *initial manipulation*. Again, it is stipulated that all compatibilist conditions for moral responsibility are satisfied, and there seems to be nothing to rule out that stipulation. Again, Pereboom thinks it is fairly plain that Plum<sub>2</sub> is not responsible. Matters are less clear in the next case.

Case 3. Plum is an ordinary human being, except that he was determined by the rigorous training practices of his home and community so that he is often but not exclusively rationally egoistic (exactly as egoistic as in Cases 1 and 2). His training took place at too early an age for him to have had the ability to prevent or alter the practices that determined his character. In his current circumstances, Plum is thereby caused to undertake the moderately reasons-responsive process and to possess the first- and second-order desires that result in his killing White. He has the general ability to grasp, apply, and regulate his behavior by moral reasons, but in these circumstances, the egoistic reasons are very powerful, and hence the rigorous training practices of his upbringing deterministically result in his act of murder. Nevertheless, he does not act because of an irresistible desire. (2001: 114)

At this point Pereboom begins to argue on the basis of ‘no sufficient difference’ reasoning. He asks what the difference could be between Plum<sub>3</sub> and Plum<sub>2</sub> that could get Plum<sub>3</sub> off the hook for killing Ms. White. There doesn’t seem to be, he says, anything that can do this work. Plum<sub>2</sub> was a case of neuroscientific preprogramming, and Plum<sub>3</sub> is a case of social preprogramming. Why should that

difference in causal means make a difference to the subject's moral accountability? Plum<sub>3</sub> is a case of *social indoctrination*.

Case 4. Physicalist determinism is true, and Plum is an ordinary human being, generated and raised under normal circumstances, who is often but not exclusively rationally egoistic (exactly as egoistic as in Cases 1–3). Plum's killing of White comes about as a result of his undertaking the moderately reasons-responsive process of deliberation, he exhibits the specified organization of first- and second-order desires, and he does not act because of an irresistible desire. He has the general ability to grasp, apply, and regulate his behavior by moral reasons, but in these circumstances the egoistic reasons are very powerful, and together with background circumstances they deterministically result in his act of murder. (2001: 115)

With Plum<sub>4</sub> we reach a scenario very close to our own case (assuming determinism is true). It is a case of *regular determined decision-making*. But Pereboom presses that point that the compatibilist cannot claim that Plum<sub>4</sub> is not responsible. Where, then, does he jump off in the sequence? There is no point at which a relevant difference obtains; 'between each successive pair of cases there is no divergence at all in factors that could plausibly make a difference for moral responsibility' (2001: 116) says Pereboom. He summarises:

The best explanation for the intuition that Plum is not morally responsible in the first three cases is that his action results from a deterministic causal process that traces back to factors beyond his control. Because Plum is also causally determined in this way in Case 4, we should conclude that here too Plum is not morally responsible for the same reason. More generally, if an action results from a deterministic causal process that traces back to factors beyond the agent's control, then he is not morally responsible for it. (2001: 116)

Pereboom is certainly right that a deterministic causal chain tracing back to factors beyond the agent's control is *an* explanation for why it is that the early Plums are not responsible. It is an explanation that the compatibilist cannot use. But the consequence-argument-style piece of reasoning Pereboom gives is not necessary, I think. It is enough for Pereboom to say that there is *something* that makes the early Plums not responsible, and the incompatibilist can provide a ready explanation of that fact: it is because determinism precludes moral responsibility. It might not be obvious in ordinary cases, but manipulation cases help bring it to light what a problem determinism really is.

We can summarise Pereboom's four cases as follows:

Case 1: Ongoing Invasive Manipulation

Case 2 (Zygote case): Initial Invasive Manipulation

Case 3: Initial Social Manipulation

Case 4: Regular Determined Action

As one can see from the parenthetical insertion, I take Pereboom's second case to be identical in the essentials to Mele's zygote case. In both cases it is the creation of one individual by another individual or body of individuals with the express purpose of exploiting deterministic knowledge to produce certain actions from the created subject at a future date. Mele's zygote argument is therefore basically an argument from the obvious failure of responsibility in Case 2 to the failure of moral responsibility in a deterministic universe more generally. But Pereboom's four-case presentation is, I'd say, credibly considered to add up to a slightly punchier manipulation argument than Mele's. The step-by-step presentation of small adjustments Pereboom gives makes the positing of an abrupt change from non-responsibility to responsibility look a more tendentious and implausible affair.

I describe the first two cases as 'invasive'. There seems a clear sense in which this is so. The neuroscientists in the first two cases are directly fiddling with Plum's brain. They have bypassed the usual methods we employ to change someone's mind or attitude about something (argument, persuasion, encouragement, pleasing presentations, and so on), and have brought about these states directly. It is in that sense that I consider the first two cases invasive. The third case is not invasive in that way. There, the manipulation occurs through channels we are more familiar with: the indoctrination of a youth through strict training and instruction, a social matter.

Lastly, some confusion has arisen because of the way that Pereboom described Case 1. The expression 'directly producing his [Plum<sub>1</sub>'s] every state from moment to moment' has led Demetriou (2010) to claim Plum<sub>1</sub> is indeed not responsible, but not because of anything to do with manipulation. He is not responsible because he doesn't satisfy basic conditions of agency: his later psychological states aren't caused by his earlier ones, they are caused by the neuroscientists at every instant. If each one of my mental states is caused by something external to me at every moment, and not by my prior states, in what sense, suggests Demetriou, can I be a functioning agent? Don't intentions have to be caused by earlier psychological states (or some other states concerned with the agent) to be considered the genuine acts of the subject? Pereboom adjusted recent presentations of his argument (2014) in response to her criticism. But I believe Matheson is correct in dismissing these worries and sticking to the original form of the argument. He points out that 'It is not as if

normally brain/mental states alone are sufficient to produce subsequent brain/mental states; there are always background conditions which are jointly sufficient with the prior brain/mental state.’ (2016: 1976). Thus, it isn’t true that the neuroscientists’ influence must suppress the prior brain/mental states of Plum<sub>1</sub>, because they, like the external world, for instance, can be part of the total setup that causally influences any one of Plum<sub>1</sub>’s later states. Matheson also makes the point that even if a somewhat invasive ‘first intervention’ is necessary to alter Plum<sub>1</sub> psychologically so that he comes to decide to kill Ms. White, after that the neuroscientists role can be reduced to that of mere ‘causal enforcer’: ‘The neuroscientists would still be manipulating Plum<sub>1</sub> “moment to moment” because they are still controlling exactly what his states are like from moment to moment; it’s just that they have less work to do after the first intervention, which does much of the heavy lifting.’ (2016: 1975). Thus, even if some suppression of Plum<sub>1</sub>’s prior mental states is necessary originally, after that only minor adjustments might be needed to keep him on track—minor adjustments which don’t suppress prior psychological states. So, as long as we understand Case 1’s ‘directly producing his every state from moment to moment’ in the manner Matheson proposes, I see no reason to think that Plum<sub>1</sub> fails basic conditions of agency.

#### **5.4 The Hard-Line Soft-Line Distinction**

I now turn to more general criticisms and objections to the Manipulation argument. But one important distinction lies in the type of reply that one should give to the manipulation argument. McKenna (2008) distinguishes between hard-line and soft-line replies. Suppose we present the manipulation argument as follows:

(a) If an agent *S* is manipulated to *A* in the manner of the Zygote case, then *S* is not morally responsible for his *A*-ing.

(b) An agent *S*’s being manipulated to *A* in the manner of the Zygote case is not relevantly different from the actions of regular determined agents satisfying compatibilist conditions for moral responsibility.

(c) Ergo, regular determined agents satisfying compatibilist conditions for moral responsibility are not morally responsible for their actions.

A hard-liner will deny (a). A soft-liner will deny (b). Soft-line replies have been proposed by e.g. Fischer (2004) and Demetriou (2010). Such replies try to locate some relevant difference between

the manipulated agent and the agent acting in regular determined conditions. For my part I think that objecting to details of the manipulated case (in the manner that Demetriou did with Plum<sub>1</sub>, say) and insisting that that detail is relevant and doesn't carry over to regular determined cases is an ultimately futile endeavour. As McKenna puts it:

[A]s I see it, for any world at which determinism is true, it is at least in principle a metaphysical possibility that, whatever the causal springs of an agent's actions happen to be, there is a way to replicate them by artificial means. So, eventually, incompatibilists will be able to come up with some manipulation story that gets the causal details "just right" as would be required by whatever credible formulation of CAS (Compatibilist-friendly Agential Structure) anyone could cook up. (2012: 171)

I think he is exactly right. The reply compatibilists make to the manipulation argument therefore has to be a hard-line reply.

McKenna (2008) defends such a hard-line reply himself. He defends what he calls the 'reverse generalisation' strategy. To see how that reply works, consider Pereboom's four-case presentation. The thought behind Pereboom's presentation is that you judge Plum<sub>1</sub> to be non-responsible, and because Plum<sub>2</sub> doesn't appear relevantly different, you carry that judgement over to Plum<sub>2</sub>. And likewise you carry that judgement from Plum<sub>2</sub>'s case over to Plum<sub>3</sub>, and then to Plum<sub>4</sub>. But then you have reached an incompatibilist conclusion. McKenna proposes that the compatibilist is within his rights to simply run the argument the other way. The compatibilist should start with Plum<sub>4</sub>, where he has no qualms about stating that Plum<sub>4</sub> is perfectly morally responsible, and then, perceiving that there is no responsibility-relevant difference between that case and Plum<sub>3</sub>, he likewise claims that there is no issue with saying that Plum<sub>3</sub> is morally responsible, and thereby justifies moving all the way to affirming that he has no reason to think Plum<sub>1</sub> isn't morally responsible. The force of the argument all seems to depend where you start from. If you start with Plum<sub>1</sub> and move up, then you end up an incompatibilist; if you start with Plum<sub>4</sub> and move back, then you end up a compatibilist. But if an argument can be employed equally well by either side, then the argument supports neither side, and thus cannot be used as evidence for incompatibilism.

Fischer has also, in effect, argued in the same way. He writes,

[I]f they show anything, [manipulation arguments show that] there is no difference between certain "initial design" scenarios and ordinary scenarios in which there is no special reason to doubt compatibilism. If this is correct, then I do not see how the Zygote Argument and its siblings show that the price of compatibilism is exorbitant – or even marginally higher than it has always been.

How could it increase the cost of compatibilism to show that a compatibilist must accept that an agent is morally responsible in a scenario that is no different than an ordinary situation in which there is no special reason to call into question the agent's moral responsibility? (2011: 5)

The same strategy is present here. Start with the ordinary case of determination. There is no evident problem with compatibilism there. Fine; so, since the manipulated case is not relevantly different, there is no problem there either.

But I think that such a bare 'reverse generalisation' strategy is an inadequate response to the manipulation argument, and does nothing (or next to nothing) to blunt the intuitive force of the manipulation argument. I think there are two points that can help us see this. First, Fischer accepts that in the case of regular determination, we are to characterise the situation as one in which 'there is no special reason to call into question the agent's moral responsibility'. In effect, a case of agnosticism in the regular case. The attitude is, 'Well, I don't see any reason to *doubt* that they are responsible. I don't see directly see that they *are* responsible either, but I don't have any reason to doubt it.' Determinism isn't obviously relevant to moral responsibility, nor is it obviously irrelevant. Arguments must be given to motivate the idea that determination threatens moral responsibility. Mere intuitive reflection on an instance of determined decision-making won't tell us anything either way. But then if we are, in the first instance, and relative to the initial data of intuition, agnostic about whether being determined is compatible with moral responsibility, it isn't legitimate to carry over that agnosticism when we encounter a case that would move us away from our indifference. This is because our agnosticism obtains precisely because we can see no reason to prefer one view over the other. But when such a reason is encountered, our views should shift accordingly. We began by thinking we had no reason, on the basis of the ordinary, regular case of determined decision-making, to doubt that the compatibilist conditions were sufficient for moral responsibility. When we think about the manipulation case, however, we feel it is wrong to blame the victim of manipulation for wrong acts they were manipulated into. And we thereby acquire a reason to doubt that the proposed compatibilist conditions are sufficient.

This is surely a regular occurrence in philosophy. A young student is so impressed with the horror of execution that he instinctively believes that the death penalty is impermissible. But then someone asks him to consider the case of a brutal mass-murderer, and all of a sudden the death penalty seems permissible now in certain cases. Perhaps the example is ultimately unpersuasive, but that's not the point. Having a plausible general claim overturned on the basis of a hard-to-initially-think-of but intuitive counterexample is a staple of philosophical debate and progress. Indeed, Fischer himself finds the Frankfurt counterexamples persuasive. Suppose someone responds to the

Frankfurt counterexample by saying that Smith can't be responsible for killing Jones because 'Smith couldn't do otherwise thanks to Black's device, and you can't be responsible when you can't do otherwise!'. Logic doesn't prevent such a move, but most people would take the intuitive cost of denying that Smith is responsible to be too great for that move to be feasible. Likewise, the cost of taking manipulated agents to be morally responsible will be held by many to be too great.

Secondly, we can see more clearly what is wrong with the 'reverse generalisation' response by considering matters in terms of sets of beliefs. McKenna's response to Pereboom's four-case argument was to suggest that one can just as easily reverse the argument, and that it really all hinges on where you start from, for it is your starting position that determines your entailments. But that isn't really true. Rather, it all hinges on the *comparative collective plausibility* of your position and its entailments. Suppose we do, as McKenna suggests, start with Plum<sub>4</sub> and work our way back. If we do that, then we infer that Plum<sub>3</sub> is responsible, Plum<sub>2</sub> is responsible, and also Plum<sub>1</sub>. So our set of beliefs looks like this:

B<sub>1</sub>: {Plum<sub>4</sub> is responsible, Plum<sub>3</sub> is responsible, Plum<sub>2</sub> is responsible, Plum<sub>1</sub> is responsible}

And if we work our way forwards from Plum<sub>1</sub>, then our set of beliefs looks like this:

B<sub>2</sub>: {Plum<sub>1</sub> is not responsible, Plum<sub>2</sub> is not responsible, Plum<sub>3</sub> is not responsible, Plum<sub>4</sub> is not responsible}

If we start with the belief that Plum<sub>4</sub> is responsible, then we end up with B<sub>1</sub>. If we start with the belief that Plum<sub>1</sub> is not responsible, then we end up with B<sub>2</sub>. The two sets of beliefs collect the relevant entailments. But that is not where matters rest. We must ask, which set of beliefs is the most plausible? Which set is more difficult, intuitively, to believe? But when we ask that question, it becomes obvious that B<sub>1</sub> is more implausible, all elements considered, than B<sub>2</sub>. B<sub>1</sub> commits one to thinking that Plum<sub>1</sub> and Plum<sub>2</sub> are morally responsible. That is very counterintuitive. Does B<sub>2</sub> contain anything comparably counterintuitive? It does not. One might try to push the thought that <Plum<sub>4</sub> is not responsible> is counterintuitive. But this seems a stretch. People are intuitively unsure about whether regular cases of determined decision-making are responsibility-cancelling. The average Joe will be unsure. If anything, the intuitive worries the layperson has about determinism will incline him more to the view that Plum<sub>4</sub> is indeed not responsible. B<sub>1</sub> is therefore evidently the more implausible set of beliefs. But B<sub>1</sub> is the compatibilist's set of beliefs; B<sub>2</sub> is the incompatibilist's set. An agnostic undecided on the two views is therefore put under rational

pressure to prefer incompatibilism. But then the manipulation argument succeeds: contemplation of these manipulated cases puts pressure on one to prefer incompatibilism. More clout must therefore be given to hard-line replies than merely the suggestion that one might reverse the generalisation. (I will offer such clout later in this chapter.)

### 5.5 Is Manipulation is the Problem?

One important category of soft-line reply is to suggest that manipulation itself is the problem. So, if we recall the structure of the manipulation argument,

(a) If an agent *S* is manipulated to *A* in the manner of the Zygote case, then *S* is not morally responsible for his *A*-ing.

(b) An agent *S*'s being manipulated to *A* in the manner of the Zygote case is not relevantly different from the actions of regular determined agents satisfying compatibilist conditions for moral responsibility.

(c) Ergo, regular determined agents satisfying compatibilist conditions for moral responsibility are not morally responsible for their actions,

one response is to deny (b), and locate the relevant difference in precisely the fact that in the zygote case the agent is manipulated and in regular cases the determined agent is not. King (2013) has complained that there is no feature of manipulated cases that can be plausibly held to generalise across from manipulated cases to determined but non-manipulated ones. The intuition of non-responsibility is bound-up with manipulation and doesn't survive in cases where the manipulator is taken away and determinism preserved. A compatibilist might therefore suggest that being determined via manipulation removes moral responsibility but being determined by nature (or however determinism would go if it were true in the actual world) does not. All we have to do is add a 'no manipulation by other agents' clause to our moral responsibility conditions. Lycan appears to suggest this (1987).

In response, three points can be made. First, such a response is of no use to the (compatibilist) theological determinist. This is because all determination is by God is determination by an agent—God, and therefore manipulation. Theological determination implies a global, all-encompassing manipulation. All our regular decision-making is therefore relevantly like the zygote case, if theological determinism is true.

Secondly, it isn't a plausible addition to make. It is *ad hoc* and implausible. It is *ad hoc* because there is no independent motivation for the extra condition—it sticks out like a sore thumb. That it is implausible can be seen from thought experiments like the following. Consider your life on the assumption that determinism is true. If you are a compatibilist, you will have no problem with holding yourself responsible for it. Now consider a world in which, again, determinism is true, and your life is fixed in every detail, but this time there is a manipulator at the beginning of your life who has programmed you to perform, eventually, every action they wanted. If the intrinsic features of your life are kept identical, along with every experience, it is odd to suppose that moral responsibility obtains straightforwardly in one case and not in the other. It is also implausible to take a 'no manipulation by other agents' condition as a basic condition. Surely manipulation is bad because it violates a more fundamental condition. But if that is so, then that condition should be spelled out.

Thirdly, Pereboom provides a good counter to this suggestion himself:

One distinguishing feature of Case 4 is that the causal determination of Plum's crime is not, in the last analysis, brought about by other agents. However, the claim that this is a relevant difference is implausible. Imagine a further case that is exactly the same as, say, Case 1 or Case 2, except that Plum's states are induced by a machine that is generated spontaneously, without intelligent design. Would he then be morally responsible? (2001: 115)

Pereboom appears right in saying that if, in every manipulation case, we replace the neuroscientist or other agents with mere machines instead, that the intuition of non-responsibility yet remains.

## 5.6 Bignon's Response

Bignon is a theological determinist, and in his recent monograph he offers a response to the manipulation argument from a theological perspective. Here is the solution he proposes that the theological determinist should employ to counter the manipulation argument:

Let me then suggest that the relevant difference we are looking for is predicated upon a moral principle much like the following: "in order for a human choice to be morally responsible, it is necessary that the choice be made on the basis of that person's *God-given character and desires*." In other words, for a choice to be free such that its maker is morally responsible, it need not be

undetermined, but it does need to be determined (assuming determinism is true) by the agent's own desires, which flow from the agent's God-given character and inclinations. (2018: 36–37, italics his)

This distinctively theological suggestion allows Bignon to explain why it is that people in manipulated cases are not responsible, while those in the regular determined cases are: when neuroscientists or demi-gods are interfering with people's psychology to bring about the decisions the neuroscientists or demi-gods want, then it becomes false that the agent thus manipulated is acting on their God-given character and desires. The character and desires have now been given (at least in part) by the manipulating party.

Despite our shared commitment to theological determinism, I do not find Bignon's solution satisfactory. For one thing, it appears to be solution by fiat where divine manipulation is concerned. Theological determinism faces a harder problem of manipulation than regular deterministic compatibilists because manipulation is everywhere for the theological determinist: God is manipulating everything any human being does. That seems like a problem, but Bignon asserts that the problem is the solution! If God is giving you your character and desires, then that is alright; that's when you are morally responsible. But that is surely too quick. Those who find manipulation cases intuitively problematic will likewise find Bignon's suggestion intuitively problematic, because Bignon's suggestion amounts to the claim that when God does the manipulation responsibility obtains (or at least there is no issue on that score with responsibility obtaining).

For another thing, it seems *ad hoc* in the same way that the suggestion that it is bare manipulation that is the problem was *ad hoc*. Indeed, it is really only a refined version of it. The suggestion that 'manipulation is the problem' has become 'non-divine manipulation is the problem'. And to the extent that it is more refined, it is even more *ad hoc*. Why should somebody's being given their character and desires by God be consistent with responsibility while those desires and character being given in the same way, but by a different agent, a human agent, suddenly become inconsistent with it? And even if God as manipulator makes a difference, shouldn't this difference be explicable in more fundamental terms? Then what are those more fundamental notions? And, again, it is implausible to imagine that your life if caused by God would raise no worries for moral responsibility, but an intrinsic duplicate of your life but caused by a human or demi-god manipulator would be inconsistent with it. And isn't everything given by God on theological determinism? Why don't human manipulators count as God's means of giving characters and desires to the manipulated subjects? For all these reasons, I think we need more than Bignon provides.

## 5.7 The Uncertainty Response

I have thus far pooh-poohed the existing responses to the manipulation argument, both secular and theological. But I do think that a satisfactory response to it can be given, and I shall do that here in the two closing sections. In this section I shall give a relatively weak response, granting much of the force of the manipulation argument, but putting a slightly different spin on matters. In the last section I give a more forceful, hard-line reply to the manipulation argument, arguing that the intuitive push manipulation cases exert against compatibilism can be matched by an equally intuitive push against incompatibilism.

I introduce the first of these two responses, the Uncertainty Response, by way of Patrick Todd's clever suggestion concerning manipulation arguments. Todd says this:

Traditional manipulation arguments present cases in which manipulated agents meet all compatibilist conditions for moral responsibility, but are (allegedly) not responsible for their behavior. I argue, however, that incompatibilists can make do with the more modest (and harder to resist) claim that the manipulation in question is *mitigating* with respect to moral responsibility. The focus solely on whether a manipulated agent *is* or *is not* morally responsible has, I believe, masked the full force of manipulation-style arguments against compatibilism. (2011: 128, italics his)

Recall the difference between Plum<sub>2</sub> and Plum<sub>4</sub>. Todd says that the advocate of the manipulation argument doesn't need to push for the admission, from the compatibilist, that Plum<sub>2</sub> is not responsible. He needs only to push for the gentler admission that Plum<sub>2</sub> is *less* blameworthy than Plum<sub>4</sub>. Surely, argues Todd, it is intuitive that Plum<sub>2</sub> is not as blameworthy as Plum<sub>4</sub>; after all, Plum<sub>2</sub> was manipulated. On a blameworthiness scale of 1 to 10, would you, O compatibilist, even if you think both are responsible, rank both Plum<sub>2</sub> and Plum<sub>4</sub> as equally responsible, both as a 10? Surely not.

But once it is conceded by the compatibilist that Plum<sub>2</sub> isn't as blameworthy as a regular determined agent, then it looks like the compatibilist is committed to thinking that determination is relevant to moral responsibility: it *mitigates* it. Todd continues:

But if the compatibilist admits that determinism itself is mitigating, a fair question is, In virtue of what? What is it about determinism's obtaining that makes revised judgments of blameworthiness appropriate? Here the compatibilist is on thin ice, for she must specify features of determinism that

only *mitigate* responsibility rather than *ruling it out*. Now, what could such features be? I submit that I cannot see what the compatibilist could offer here. (2011: 131, italics his)

Suppose that the compatibilist says that manipulating determining mitigates responsibility because it implies our characters are formed ultimately by factors beyond our control. But that holds universally on determinism. So that wouldn't mitigate, it would rule out responsibility entirely. Suppose the compatibilist says that manipulating determining mitigates responsibility because it undermines the ability to do otherwise. But if manipulating determining does that, so will regular determining. In short, Todd argues, there is no plausible way of explaining why Plum<sub>2</sub> is of reduced blame save by appeal to determinism *simpliciter*. But if determinism diminishes moral responsibility, then it does so across the board. And that looks like a problem for compatibilism.

But I think that Todd has overlooked an important distinction between the two following propositions:

(I) I judge that Plum<sub>2</sub> is less blameworthy

and

(II) I am less confident about my judgement that Plum<sub>2</sub> is blameworthy.

To admit (I) does look like it raises thorny issues for the compatibilist—it commits one to a belief in lesser responsibility in Plum<sub>2</sub>'s case. But (II) does *not*. It simply acknowledges that matters are more complicated and less certain here, and even though Plum<sub>2</sub> might be fully blameworthy, that is harder to see than in the regular case. In other words, if (II) is true, then the degrees obtain this side of the mind/world divide, as an epistemic feature, not a metaphysical feature.

Indeed, I would like to suggest that that is a perfectly viable way of understanding the manipulation cases. Plum<sub>1</sub>, Plum<sub>2</sub> and Ernie in the Zygote case are all cases where we are *uncertain* whether or not they are responsible. This is what our reluctance to rank the blameworthiness of Plum<sub>2</sub> and high as Plum<sub>4</sub> amounts to: an uncertainty about the moral responsibility of Plum<sub>2</sub>. It's a very unusual scenario, and the compatibilist just isn't sure what to make of it. No admission of lesser responsibility need follow.

Now, this isn't sufficient as a compatibilist rebuttal of the manipulation argument, because if it is conceded that uncertainty reigns as far as Plum<sub>2</sub> is concerned, then the incompatibilist can push the point and argue that, because Plum<sub>4</sub> is principally no different from Plum<sub>2</sub>, then the uncertainty must carry over. The compatibilist taking this line of response must be just as uncertain about Plum<sub>4</sub>'s responsibility as he is about Plum<sub>2</sub>'s.

And I would concede this point. Or, rather, I would concede that *other things being equal* the compatibilist taking the uncertainty response should indeed be uncertain and agnostic about responsibility in more commonplace non-manipulative cases of determination. But here the theological determinist has an advantage his secular counterpart does not. Other things are *not* equal for the theological determinist. The theological determinist does not think that the empirical and intuitive data we have about our own responsibility in cases of regular determination exhaust everything of evidential relevance. The theological determinist (when Calvinist) will also take himself to have good reason to believe that the Scriptures are divine revelation. But the Scriptures assert, or rather presuppose, man's responsibility in the ordinary cases of everyday life. Thus, if the theological determinist takes himself to have good reason to believe from the Scriptures that (i) man is responsible in everyday cases, and from the Scriptures and elsewhere that (ii) theological determinism is true, then he will have access to reasons that will enable him to overcome the uncertainty judgement about moral responsibility in ordinary determined cases. That is the uncertainty response to the manipulation argument.

## **5.8 A Manipulation Argument for Compatibilism**

The uncertainty response is a broadly defensive one. But here is a more aggressive pushback against the incompatibilist advocate of the manipulation argument. I will introduce the argument by telling a story:

You are Jack. You have a good friend called 'Algernon'. You and Algy went to the same school together, and both entered the world of London business where you both find success and offer each other valuable assistance in your various commercial ventures. And one day scientists make a striking discovery. They discover that there existed, millennia ago, a certain god called 'Diana'. Not only have they discovered her existence, they have also discovered that she has interfered in a remarkable way with about a tenth of the human population. Through interfering with human zygotes at the early stages in human history, she inserted DNA-based instructions of her own creation into the human genome. The presence of these genetic instructions enables her to effectively program every action that the grown-up human being so interfered with performs. And she exploits this to, indeed, program every action of every human being's life in which these instructions are found. Not merely that, but she has inserted instructions for all the offspring that those with the new genetic material have. Let a human being with this new genetic programming be a D-human. Diana has set things up so that when a D-human mates with another D-human, the

instructions are triggered in the child of that union, and the child also becomes a D-human, and performs whatever actions in response to stimuli that the instructions bequeath to them. In this way, D-humans remain about a tenth of the total human population. These scientists have therefore discovered that about 1 in 10 people one knows are actually acting in response to a very clever piece of genetic programming.

The scientists have also discovered that Diana has recently perished in a war of the gods. But that is of little consolation, for her effects live on, and the genetic instructions appear here to stay. But about those instructions: they do not *compel* the D-human to act. They do not override his desires and personality. Rather, they work *via* the D-human's personality, character, and so on. The instructions create the right character, the right desires, and the D-human then voluntarily (in a manner that satisfies compatibilist conditions on freedom) acts on them. The instructions cause action, but they do not compel it. A compatibilist could look at how a D-human acts and be perfectly content in ascribing moral responsibility to them.

These scientists do one last thing. They feel the public should know about these D-humans in our midst, and the scientists know how to detect the D-strain. So they have everyone's DNA scanned. When they detect the D-strain, they note down the name of the afflicted subject, and they go on to publish, in everyone's local newspaper, the names and details of every D-human in that neighbourhood.

Thus, one day, as you are enjoying your morning coffee, you open up the newspaper and are confronted with the long list of names, names of all the D-humans local to you. You know you aren't one—the scientists told you that after you got scanned—but you are worried about your acquaintances. Thankfully, you don't notice the name of anyone you know, save for one sad exception: you notice, with dismay, Algy's name is on the list. 'Algernon D. Whittaker' it reads, as plain as day. There can be no mistake—it's Algy alright.

Stirred by the news, you make your way over to Algy's manor, intending to offer him in person, like any decent chap would, your commiserations. Algy greets you in his usual warm and friendly manner, and you both take up seats in the drawing room.

'Nasty business this,' you say to to him. 'Imagine finding out that, all this time, one has actually been following instructions from a deceased deity! But, Algy, my dear fellow, I want to know how you are taking the news. How do you feel about the whole thing?'

'Well, Jack,' he replies, 'At first I *was* rather cut up about it. It felt like a sort of loss of control. Was it really me acting all these years? Who am I, really, if in everything I've ever done I was only following instructions programmed into me by another person? So, questions like that. But then I played the argument out. And then, after prolonged reflection, my attitude shifted quite radically.'

He stands up, smiling.

‘What do you mean?’ you ask.

‘I mean *this!*’ he cries, and, all of a sudden, deals you a cruel blow across the face. You are utterly shocked and bewildered. But you soon recover.

‘What in blazes was that for?’ you roar.

‘Think about it, Jack!’ he responds. ‘Everything I’ve done and will do I do because I have been manipulated by that old bird of a deity. At first I wondered if there was some way I could defeat her programming—outwit her somehow. But it’s not possible—her programming is comprehensive. I realised that, even if I tried to double-guess myself, then I was only trying that because I had been programmed to try that. There’s no escape from it.

‘But then it dawned on me what a liberating fact that is!’ Algy goes on. ‘If no matter what I do, I am following manipulative programming, then I can not be responsible for anything I do! I, and all the other D-humans, must now be considered entirely exempt from the realm of moral accountability. We can do anything we want now, without fear of being wrongdoers! The sky’s the limit! Of course, my life is no different, from the inside, than yours. I am not subject to strange compulsions (or no more than an ordinary human being is); I have to weigh and deliberate matters just as you do, but with the great and profound difference that I can not be blamed for any of my decisions! Fantastic! They can all be traced back to Diana’s manipulation (bless her!), thereby entirely exonerating me.’

With that, he deals you another cruel blow across the face. It is not to be borne. You lunge at him, tackling him to ground, and begin to pummel him with your fists. Algy is an able wrestler, however, and through a combination of lithe movement and deft parrying, he manages to deflect the force of your blows. And his mocking cry rises above the fray:

‘You still don’t get it, do you, Jack? You’re hitting an innocent man! Ha ha! Are you outraged at me, Jack? Are you indignant? Entirely inappropriate responses! I’m not a morally responsible agent! In fact, all you do by hitting me is add to your own guilt! Ha ha! I can no more be fairly punished than a robot, or a painted vase!’

There is the story. I believe it points to an intuitive cost to the incompatibilist position. The incompatibilist, it appears, is forced to think that Algernon is in no way blameworthy for his behaviour. But that seems like quite the cost. Most of us, I wager, would feel that Jack was quite within his rights to give Algy a good thrashing, Diana or no Diana. After all, from the inside, Algernon’s life was in no principal way different from yours or mine. We might also turn the question on ourselves: suppose you or I discovered that we were a D-human. Should we cease to view ourselves as responsible for anything we have done? Intuitively not, I think.

But then the compatibilist has his own intuition to appeal to to match the intuition the incompatibilist appeals to in the case of Plum<sub>2</sub> and Ernie. This vindicates a hard-line reply to the manipulation argument. If we add the proposition <Algy is not responsible for his behaviour> to the collection of propositions that the incompatibilist is committed to, then all of a sudden the compatibilist commitments in this regard do not appear clearly more counter-intuitive than the incompatibilist ones.

This latter response can be employed by the secular compatibilist too. But between them, both the uncertainty response and this manipulation argument for compatibilism constitute a sufficient rebuttal to the manipulation argument as directed at the theological determinist.

# PART 3—THEOLOGICAL DETERMINISM AND GOD'S GOODNESS

## Chapter 6—The Justice of Hell

### 6.1 Introduction

We saw that in ch. 3 that Channing appeared concerned about the justice of Hell itself, and given that so often the complaint against Calvinism concerns God's predestining to damnation, I doubt that a fully satisfactory response to the difficulty can be given unless, indeed, the question of the *justice* of Hell is also addressed. Consider that there are a few arguments one could make against both of the following claims being true: <God exists> and <Hell exists and will be occupied by a substantial number of human beings>. One big argument is that God would surely use his omnipotent power to prevent Hell from being occupied—there is no justifying good great enough that would permit God to let people end up there. That argument is the subject of the next chapter. Another popular objection relies on a perceived disparity between the gravity of human offences and the severity of the divine punishment. Isn't Hell too severe a punishment? It is that question that is addressed in this chapter.

Here, for instance, are Isaac Asimov's remarks on the matter:

I would [...] want a God who would not allow a Hell. Infinite torture can only be a punishment for infinite evil, and I don't believe that infinite evil can be said to exist even in the case of a Hitler. (1995: 334)

It is a natural thought. Hell is supposed to perform a retributive function—it is supposed to be *just*. But it is not just to punish someone in a manner clearly in excess of what they deserve. But isn't that what is going on with the case of Hell, as traditionally understood? Aren't human sins, for all their grievousness, of a finite quantity? Isn't Hell, as traditionally understood, punishment of an infinite nature? How, then, can Hell be just? The punishment seems far in excess of the crime. This

objection relies on a perceived disproportionality between the gravity of the human offence and the seriousness of the punishment of Hell. Satisfying some such ‘proportionality principle’ is therefore a requirement.

But, as we will find in the course of this chapter, this ‘proportionality principle’, strikingly, does not have the consequence that finite sin cannot merit infinite punishment. Moreover, I will outline several ways in which human sin, and human beings, can be correctly considered as being infinitely heinous. In this way, it will be shown that the charge that Hell is unjust cannot be made to stick. We will also see that the Calvinist has certain advantages in this regard over the Arminian.

## 6.2 What is Hell?

But, first, we must define our terms. What, exactly, are we to understand Hell as? For the purposes of this chapter, I shall, building on the list of theses that Jonathan Kvanvig (1993: 25) takes to characterise what he calls the ‘Strong View of Hell’, take Hell to be correctly described by the following theses:

(H1) **The Anti-Universalism Thesis.** This is the claim that God will, on the Day of Judgment, consign some people to Hell. Not everyone goes to Heaven.

(H2) **The Existence Thesis.** People consigned to Hell are not annihilated; they continue to exist. This claim, in conjunction with (H1), means that the quip sometimes heard from universalists —“I do believe in Hell; it’s just that no one will be there.”—is ruled out.

(H3) **The No-Escape Thesis.** Once consigned to Hell, there is no means of removing oneself out of Hell, nor any possibility of this happening to one, nor of it being true at any later point that one is not in Hell. This thesis therefore implies that Hell is of everlasting duration.

(H4) **The Retribution Thesis.** Hell is there to satisfy the demands of justice. The nature of one’s stay in Hell is so constituted that it metes out deserved punishment for one’s misdeeds.

While this is all good as far as it goes, (H1)–(H4) are not sufficient, I believe, to capture everything pertinent to the traditional view of Hell. This is because (H1)–(H4) do not entail that there is any significant degree of conscious suffering in Hell. It is consistent with the above theses that Hell contains only mild discomfiture, or that it consists merely in the deprivation of certain goods or pleasures, while many other goods and pleasures yet remain available. Perhaps that is all the punishment people deserve for their sins.

Consider how Dante describes those who occupy the first circle of Hell, Limbo. It contains virtuous pagans, such as Virgil, who, lacking baptism, or having no exposure to true religion, were unable to worship God aright. Dante says of the first circle that

There, as it seemed to me from listening,  
Were lamentations none, but only sighs,  
That tremble made the everlasting air.

And this arose from sorrow without torment (*Inferno*, Canto 4, 25–28)

Virgil explains to Dante exactly how he and the other noble pagans are punished:

Lost are we and are only so far punished,  
That without hope we live on in desire. (*Inferno*, Canto 4, 41–42)

An account of Hell that contained only unhappiness of this sort—an endless sense of disquiet and unfulfilled desire that prompts continual sighing—could satisfy (H1)–(H4), but it would be a picture of Hell greatly at odds with the more accustomed imagery of fiery torment.

So, theories of Hell that satisfy (H1)–(H4) may be considered insufficiently severe. We should keep to hand, therefore, the following thesis:

(H5) **The Torment Thesis.** The punishment that is meted out in Hell is such that every occupant of Hell is subject to frequent conscious suffering of a kind that is at least comparable to the worst cases of earthly suffering.

Adding (H5) to (H1)–(H4) guarantees an account of Hell that is closer to the more severe, traditional understanding.

But it is helpful to consider (H1)–(H4) apart from (H1)–(H5). For the purposes of this chapter, let ‘Hell’ refer to any model of Hell that satisfies (H1)–(H4), and let ‘Hell+’ refer to any model of Hell that satisfies (H1)–(H5). As we shall see, it is easier to make a case against the existence of Hell+ than it is against Hell, though ultimately I shall defend the plausibility of both varieties.

## 6.3 Lewis Carroll

We began with a quote from one fiction writer, Isaac Asimov. But the argument was rather short and undeveloped. However, another famous writer of fiction, Lewis Carroll, pushes the same line of argument that Asimov did, and he fleshes out his thoughts at length.<sup>36</sup>

Carroll begins his argument with an appeal to the proportionality principle we have already noted. He states it as follows:

There is, however, *one* principle which clearly applies equally to both [human and divine justice]: we recognise that some *proportion* should be observed, between the amount of crime and the amount of punishment inflicted: for instance, we should have no hesitation in condemning as unjust the conduct of a judge who, in sentencing two criminals, had awarded the *greater* punishment to the one whose crime was clearly the *lesser* of the two. (1899: 349)

Carroll then draws out what he takes the implications of this principle to be for the traditional doctrine of Hell. Consequently, he says,

We feel intuitively that sins committed by a human being during a finite period must necessarily be *finite* in amount; while punishment continued during an infinite period must necessarily be *infinite* in amount. And we feel that such a proportion is unjust. (1899: 350)

Not only that, but Carroll believed that the consequences of setting aside these feelings are dire indeed:

To set aside this intuition, and to accept, as a just and righteous act, the infliction on human beings of infinite punishment for finite sin, is virtually the abandonment of *Conscience* as a guide in questions of Right and Wrong, and the embarking, without compass or rudder, on a boundless ocean of perplexity. (1899: 352)

It is the contention of this chapter that believing in eternal punishment for earthly sins occasions nothing so drastic as a dethroning of the conscience, or anything comparable to that. Indeed, I will

---

<sup>36</sup> Lewis Carroll, concerned that his literary reputation rested on such trifles of fancy as unpunctual white rabbits and little girls that grew and shrank with surprising rapidity, was determined to compose a sober volume of essays on theological topics, as would better befit a man of the cloth. Alas, he passed away before he could succeed in this enterprise, and only the first essay, 'Eternal Punishment', which is discussed here, was completed. See Collingwood's remarks in Carroll (1899: 344).

attempt the opposite. I will show that the justice of Hell's infinite punishment follows plausibly from our moral intuitions in conjunction with certain plausible suppositions about human beings.

In order to assess this argument, we need to have a clear statement of this crucial proportionality principle before us. However, Carroll does not offer us the most perspicuous formulation of the principle. On his phrasing, the principle holds between 'amounts': a proportion must be observed between the 'amount of crime and the amount of punishment inflicted' (1899: 349). But 'amount' is sufficiently unclear that it may lead to obviously false understandings of the principle. It must not refer to simple quantity. Suppose I commit two minor sins, and my friend commits one extremely grave sin that far outstrips my two sins in the greatness of its evil. To suggest that I should receive twice as much punishment as my friend because I sinned twice and he sinned once is clearly unjust.

We must also dismiss the idea that the proportionality principle is concerned with temporal duration. There is a clear sense in which the punishment of Hell, because it is everlasting, would be of an infinite amount—it would be of infinite duration. But the concern of the proportionality principle isn't to match up the duration of punishment with the duration of the offence. A minor sin might be indulged in for a long while, and a great sin might be performed with astonishing alacrity, but we don't hold on that account that the longer sin must receive the longer punishment. But if the proportionality relation doesn't relate the number of sins to the number of punishments, or the duration of sins to the duration of punishments, then what is it relating to what?

## 6.4 The Proportionality Principle

Most contemporary writers who have remarked on the principle have taken the proportionality requirement to be a relation between the *gravity* or *seriousness* of the wrongdoing and the *intensity* of the punishment. Kershnar, for instance, says this:

By proportionality in punishment, I mean that there is a systematic positive relation between the *seriousness* of a person's wrongdoing and the maximum severity of punishment she may be given. (2018: 43, italics mine)

Charles Seymour puts it like this:

It is unjust to punish sins disproportionately to their *seriousness*. (1998: 69, italics mine)

Kenneth Himma presents the following understanding:

[P]unishment is morally justifiable if and only if it is not excessive, given the *magnitude of the wrong*. But infinite suffering, on this line of analysis, is out of proportion to any wrong that finite human beings could commit. (2003: 61, italics mine)

I think that all these thinkers are referring to the same type of property with the italicised expressions. I shall use 'severity' to refer to that type of property. We are surely all familiar with the fact that different moral offences are different in severity. It is evident to the moral intuition. Killing a dog without good cause is wrong; killing a stranger without just cause is much worse; and killing a family member for no good reason worse still. The offences increase in severity. (Kershnar speaks of a 'severity of punishment', and that is not an unnatural use, but I will not follow him in speaking of severity of punishments. One must regiment the language.)

It is surely right that it is the severity of the wrongdoing that determines the sort of punishment that is thereby deserved, rather than the mere numerical quantity of these sins or their temporal duration. That is not to say that numerical quantity is irrelevant, however. The sort of punishment one deserves for a lifetime's sins will be a function of both the number and the severity of those sins. But I don't wish to complicate matters by bringing in quantity at this point. Let us stick with individual sins for the moment. We therefore have one half of the proportionality principle, as follows:

(Proto-PP) For any culpable fault  $F$  exemplified by an agent  $A$ , there is a punishment  $P$  that  $A$  deserves, and  $P$  is proportional to the *severity* of  $F$ .

But the reference to punishment is uninformative. In what sense can a punishment be proportionate? But again, it is evident that, just as there are various degrees of severity, so there are various degrees of punishment. Receiving 'six of the best' to one's posterior is a lot worse than a rap on the knuckles from a ruler, yet a caning pales in comparison with a serious whipping from a cat-o'-nine-tails. I shall use 'degrees' to speak of these different levels of punishment. And the concern of the proportionality principle is to have the severity of the offence proportional to the degree of punishment. We can therefore state the proportionality principle as follows:

(PP) For any culpable fault  $F$  exemplified by an agent  $A$ , there is a *degree* of punishment  $D$  that  $A$  deserves, and  $D$  is proportional to the *severity* of  $F$ .

But more needs to be said about degrees of punishment. I shall take degrees of punishment to be sets of possible punishments. There are many possible ways to be punished for a crime. One might be beaten, starved, subjected to extreme heat, to extreme cold, to psychological torment, to public humiliation, and so on. Punishments might be comparatively mild too: one might be restricted, in certain ways, in one's freedom, or one might lose certain rights, or sources of pleasure or happiness, without there being any positive infliction of pain or suffering. But some of these means of punishment will be just as bad as each other. Perhaps being eaten alive by dogs is just as bad a punishment as being burnt at the stake. In that case, a moral offence that would have one as a proportionate punishment would also have the other. In this way, a moral offence will have, proportionate to its severity, a set of possible punishments. This set of possible punishments is the degree of punishment.

We can say more about these degrees of punishment. They are surely sets of *pairs*. This is because we don't just talk about the type of punishment that is to be meted out, but also about how long the punishment should last. As Himma notes, "any given punishment has two dimensions, duration and intensity, and both dimensions necessarily play a role in determining whether a punishment is proportional to the crime." (2003: 64). A very painful punishment with a short duration might be equivalent to a less painful punishment with a longer duration. Perhaps eight minutes with one's hand in the fire is just as bad as five minutes with one's arm in the fire. A degree of punishment, then, as a set of pairs of intensity and duration, might end up looking something like this:

{<hand in the fire, 8 minutes>, <arm in the fire, 5 minutes>, <serious itch, 1 year> ... }

Each member of the set would be a suitable punishment for a moral offence severe enough to warrant that degree. We might also, if we wished, try to bring the punishments under a common metric. To borrow Jeremy Bentham's terminology, we might coin the 'dolor' as a unit of pain or suffering. This may permit us to collapse, say, <arm being sawed off, 1 minute> and <leg being sawed off, 1 minute> to <100 dolors (per second), (for) 1 minute>, for example.

But, as noted above and as will be seen below, we can't always think of punishment in terms of the infliction of suffering. So, although some punishment-pairs may be collapsed in this way, it can't be comprehensively done.

## 6.4 Can finite sins merit only finite punishment?

We can now begin to analyse the claims of Asimov and Carroll. Asimov claimed that ‘infinite torture can only be a punishment for infinite evil’ (1995: 334). But if a just infinite punishment implies infinite evil, then, by contraposition, finite evil implies the absence of any just infinite punishment. But since our sins (assuming they are finite) still surely merit punishment (at least sometimes), and the merited punishment cannot be infinite, it follows that the finite evil for which we are responsible can only merit finite punishment. Call this claim (FF):

(FF) Finite sins can only merit finite punishment.

This claim is crucial to the popular case against the justice of Hell as given by Asimov and Carroll, and I will argue that it is false, although I don’t think its falsity is as significant as might be thought. But the components of the claim need to be analysed. What does it mean for sins to be finite? And what does it mean for punishment to be infinite?

Let’s begin with sins. In what sense are our sins finite? I see two relevant ways in which our sins are plausibly finite: they are finite in *quantity* and *severity*. As to quantity, if it were revealed to us that we were beings who had lived for an eternity past before our present embodiment in these human bodies, and that our previous existences which went back throughout all of the infinite past were as consistently sinful as our present ones, then it would surely be conceded that we all have racked up an infinite quantity of sins to our account. But no one thinks anything like that is true. We have only lived a finite amount of time, and therefore only committed a finite amount of sins. One might put the argument this way: a sin is a decision, and we have only ever made a finite number of decisions; therefore, our sins are finite in number. Thus, it is likely that all would agree that our sins are finite in *quantity*.

But what about *severity*? What does it mean for severity to be finite (or infinite)? Here, it is hard to give an informative answer. It is tempting to reach for interpretations that render (FF) trivially true. If we say that ‘finite sins’ are just those sins that merit finite punishment, then (FF) becomes the not particularly enlightening proposition that <sins that merit only finite punishment can only merit finite punishment>. Needless to say, that is nothing to write home about—it is a dialectically useless claim—the believer in Hell will of course deny that our sins are finite in that sense—the objector to the justice of Hell needs some account of ‘finite sin’ such that the believer in the justice of Hell would at least be somewhat inclined to think that the sins we commit are finite in that sense.

One way to argue, from premises all would accept, that all our sins are of finite severity would be as follows: it is clear that we commit *some* sins of finite severity. To claim that all our offences are of infinite severity is too much. So if indulging feelings of impatience while waiting in a queue is a sin of finite severity, then we have a handle on what it would mean for a sin to be of infinite severity: it would be a sin that immeasurably dwarfs sins like *that*. But it seems wrong to say, even of the grosser sins that human beings commit, such as brutal murder, that they are infinitely greater than indulging feelings of impatience, substantially greater, yes, but not *infinitely* greater.

However, defenders of the justice of Hell may deny that we commit any sins of finite severity. All sins are sins against God, they may say, and for that reason are all of infinite severity. But will I discuss that sort of reasoning below. Let us grant for the moment that the claim that all human sins are of finite *severity* is well-motivated.

But what sense are we to attach to the idea of *punishment's* being finite or infinite? In our terms, this must be an infinite degree of punishment. But degrees, as I defined them, are sets of pairs of duration and intensity. There are therefore two ways in which a degree of punishment can be finite or infinite: it can be infinite by virtue of having an infinite *duration*, or by having infinite *intensity*.

As for duration, there is a plain sense in which the punishment of Hell is infinite in that way: it has an infinite temporal extension—it goes on forever. A punishment can therefore be finite by lasting for a finite duration, and infinite by lasting for an infinite duration.

But there is the distinction, that must be kept in mind, between actual and potential infinity. A potentially infinite collection of objects is a finite collection of objects that increases in number indefinitely. Such a collection may start with a single object, or none, but more and more objects will be added without the process ever stopping. Such a collection will at no point be infinite, for an infinite sum cannot be reached by successive addition; it is instead a collection of an endlessly increasing finite number of objects. An actual infinite collection, on the other hand, is a collection that is infinite as it stands—an actually infinite library of books contains an infinite number of books as it is; you don't need to add more books to it in any way to make it infinite.

In the light of that distinction, it is clear that Hell must count as potentially infinite. The fact that no one in Hell will ever be able to claim, at any point, that they have been there for an infinite amount of time shouldn't cause us to cease from describing Hell as an infinite punishment, preferring instead to call it endless, perhaps, for it is clearly worse to be sentenced to a certain punishment of endless duration than it is to be sentenced to the same punishment for any finite duration; thus endless duration possesses the property, noted above as characteristic of infinity, of being greater than any finite duration. Thus, we can attach a clear sense to punishment's being finite or infinite by virtue of its having finite or potentially infinite duration.

What about intensity? Might a punishment be infinite on account of being of infinite *intensity*? I think there is cause for scepticism here. To begin with, although it may be possible for pains to continue indefinitely in intensity, it is not so clear that it is possible for human beings, as they are currently constituted to experience pains above a certain level degree of intensity. There is only so much, it may be thought, that body or soul may possibly bear. However, it may be that the damned in Hell are physically and/or psychically reconstituted such that they become capable of experiencing pains beyond what it was possible for them to experience in this life. But, assuming that the relevant changes in nature are made, would it be possible for a human being to experience an *infinite* pain? I don't mean a pain extended over an infinite amount of time, I mean an occurrence of a pain of infinite intensity. Here some question marks may be placed over the conceptual coherence of the idea. We readily understand that pains may increase indefinitely in intensity, but that there should be such a thing as an infinite pain? This seems like mistake, a category error of some sort. I shall follow the gut sense here, and assume pain of an infinite intensity is not possible.<sup>37</sup> Accordingly, if the suffering of Hell, or any punishment at all, is going to be considered an infinite punishment, then it will have to be by virtue of its temporal extension, its duration, and not because of the intensity of the suffering experienced therein.

## 6.5 The Central Argument

We are now in a position to state the central argument of this chapter against the justice of Hell:

- (1) Hell is a perfect expression of justice: you suffer neither more nor less than you deserve in Hell.
- (2) You are punished only for the sins of your earthly lifetime in Hell.
- (PP) For any culpable fault *F* exemplified by an agent *A*, there is a *degree* of punishment *D* that *A* deserves, and *D* is proportional to the *severity* of *F*.
- (3) The sins of our earthly lifetime are all finite in severity.
- (4) The sins of our earthly lifetime are finite in quantity.
- (FF') Sins that are of both finite severity and finite quantity cannot merit a degree of punishment that is of infinite duration. (From PP)

---

<sup>37</sup> Theologically conservative Christians have extra reason to be suspicious of the possibility of an infinite pain. If there were such a thing, then it should be possible for human beings to experience it, given the right natures. But then it would be possible for human beings to pay the penalty for their sin (infinite demerit) in a finite time, and thus (i) Hell wouldn't need to be eternal, and (ii) the need for atonement is Christ wouldn't be absolute: by suffering an infinite pain on their own behalf, people could pay a satisfactory penance for their sins—a salvation by works.

(5) Hell is a punishment of infinite duration.

Therefore,

(6) Hell is not just, *contra* (1). (From 2, 3, 4, FF', 5)

(FF') is a reworking of (FF) in the light of the full statement of the Proportionality Principle (PP). That, and the argument itself, makes it clear that (FF) (or (FF')) is *not* to be identified with the proportionality principle—a mistake one might instinctively make; (FF) (or (FF')) is a *consequence* of the proportionality principle, not an analysis of it.

In this chapter I shall argue that (2), (3), (4) and (FF') are all plausibly false.

## 6.6 The Continuing-Sin Hypothesis

The denial of (2) is a very quick way of dealing with this argument. It is called the continuing-sin hypothesis, and is defended by Murray (1999). For our purposes, we can take the continuing-sin hypothesis (CSH) to be the conjunction of the two claims that

(CSH1) Human beings in Hell continue to sin

and

(CSH2) It is the truth of (CSH1), and that alone, that justifies the degree of punishment present in Hell (or Hell+).

One can see at once that this is a promising resolution of the difficulty. If human beings continue to sin in Hell, then there is no need to deny (FF), or to insist that human beings are guilty of an infinite number of sins or that sin is always of infinite severity (denials of (3) and (4)), or anything like that. The proponent of CSH can happily grant that sins are of finite severity and of finite quantity and, indeed, that human beings can only deserve punishment of a finite duration for such sins, for what makes the eternal punishment of Hell permissible for God to dole out is, he will say, that the human beings present in Hell never stop sinning. It would be quite possible for them to 'do their time' in Hell, pay the price for their sinful lives, and then be out after a few years (or however long), if they refrained from reoffending. But they never do. So, because there is never a period of time when they refrain from sinning long enough for the demands of justice to be completely satisfied with regard to them, they therefore indefinitely prolong their stay.

I think CSH is workable, especially for the Calvinist. There are a few points of concern with it that Christian philosophers might have, however.

(i) It plausibly requires a denial of incompatibilism. Arminians won't like it. For how can it be a *guaranteed* result that those in Hell will continue to sin? Surely sheer anguish over their torment will eventually prompt the damned to decide to cease from sin just so that their suffering might be over. Or, if we aren't thinking in terms of Hell+, and there is no *great* pressure to leave Hell, the simple odds of someone with libertarian free will endlessly freely choosing to reoffend is just too unlikely. But compatibilists have no problem here—God can give the damned characters (or agential natures) such that these characters always cause the damned to choose to sin rather than refrain—but such characters would be freedom-cancelling on incompatibilism. Incompatibilists insist that nothing must determine an agent's deciding, if those deciding are to be morally responsible deciding. But if there is nothing to determine that the damned will continue in sin, how can it be a guaranteed fact that they will continue to sin forever?

(ii) A second worry is that it seems to undermine the significance of the Day of Judgment. All the reprobate is really sentenced to on that great day is punishment of a finite duration, and his continual reoffending either prolongs his original sentence or requires fresh sentencing from God. But surely the sentence of the Day of Judgment is more momentous on its own account than this picture suggests. (Note that if one says that being placed in a position whereby reoffending forever is inevitable is part of what the damned are sentenced to, then that looks like an infinite punishment, and that is what gave rise to the issue in the first place.)

(iii) There is also the worry that, even if being put in a position such that perpetual sin (and therefore perpetual punishment) is inevitable is not part of the sentence *de jure*, that is nevertheless what the sentence amounts to *de facto*, and it is hard to shake the suggestion that the *de facto* aspect is morally relevant to the sentencing. But if it is relevant, then, again, it appears that perpetual sinning is part of the punishment, and that appears relevantly infinite, and nothing is gained in responding to the proportionality problem.

(iv) CSH also seems to undermine the ultimate triumph of God over sin. Justice is never finally accomplished on CSH. Sin continues forever, and God is forever responding to it punitively. There always remains fresh sin to punish, and the work of justice is never complete.

Now, it might be thought that the standard view of Hell faces the same problem. That goes on for ever, and is therefore never completed. Isn't justice likewise never accomplished on that view?

But a response can be made: CSH is committed to the continuous commission of *sins* in Hell. He who denies CSH is not. That is a plain sense in which the advocate of CSH is committed to limiting God's victory over sin, independently of issues concerning the satisfaction of justice. There

are two relevant ways in which God can be thought to be victorious over sin. The first is by causing sin to cease; the second is by punishing sin. CSH rules out the former sort of triumph while other views do not have that consequence.<sup>38</sup>

(v) The CSH rules out *reconciliationism*.<sup>39</sup> That is the view that the sort of reconciliation spoken of in Philippians 2:10–11, ‘so that at the name of Jesus every knee should bow, in heaven and on earth and under the earth, and every tongue confess that Jesus Christ is Lord, to the glory of the Father’, extends also to the damned.<sup>40</sup> As Bawulski puts it, reconciliationism’s ‘biggest distinctive, probably its *sine qua non*, is that in the eternal state all sinning ceases: God’s victory will be such so that sin shall be no more.’ (2013: 133). This does not save the damned—they still suffer in torment—but they come to recognise God’s lordship and goodness, and they honour him from Hell’s pits, and cease from sin. Their continuing torment is thus justified only on the basis of their earthly rejection of God and his Christ. It is a cost to CSH that it rules out this exegetically defended view.

Now, I don’t think these objections are decisive. Consider concern (iv), that God’s triumph over sin is limited if sin continues for ever. It can plausibly be countered that continual sinning increases God’s glory insofar as it permits God to continually display his ability to frustrate and ruin those who oppose him (a theme of the next chapter). Consider also concern (v), that CSH rules out reconciliationism. We should distinguish between hamartiological reconciliation and liturgical reconciliation. The first is a reconciliation that involves the reconciled party ceasing from sin. The second is a form of reconciliation that only involves the reconciled party participating in worship, but not necessarily refraining from sin in toto. It seems to me plausible to hold that this latter sort of reconciliation is all that is required to do justice to the reconciliation proof-texts.

Henri Blocher, a proponent of reconciliationism, writes against this idea, ‘The theory of sin forever flourishing ignores the message of Christ’s perfect victory over sin and all evil. *Every knee shall bow and every tongue confess* [...] (Phil. 2:10f), those of the lost included. It cannot mean mere outward, hypocritical and forced agreement’ (1992: 303). But one can grant that the bowing, confessing, and worshipping that the damned perform from Hell’s dark chambers is sincere and profound, whilst also holding that the damned continue to sin in other respects. After all, Christians in this life offer Christ genuine worship whilst sadly sinning in other areas.

The denial of (2) is therefore a ready manoeuvre for the Calvinist.

---

38 Perhaps the best exegetical argument that CSH is true is from the parable of Dives and Lazarus (Luke 16:19–31). Dives still considers himself to have the authority to order Lazarus about, despite the dire condition Dives finds himself in. This suggests he is still eaten up with pride. However, one might say that matters change at the eschaton, or issue a caution about taking parables too strictly.

39 For a defence of the view, see Saville 2005 and 2007, and Bawulski 2013.

40 See Col. 1:20 and 1 Cor. 15:28 for similar verses.

## 6.7 Denying FF

But one might also consider the denial of (FF'). (FF') was the claim that sins that are both finite in severity and in quantity are not able to merit an infinite degree of punishment. I think (FF') can be shown to be false, but it will turn out that its falsity is of limited use. In particular, not much use in showing that Hell+ is just.

### 6.7.1 More detail

But first I want to flesh out in more detail how to model claims like (FF'). We saw that (FF) turned out to be a principle (FF') relating moral severity to degrees of punishment. Consider, therefore, the set, S, of moral severities, containing every possible strength of severity that could, in principle, be attributed to an act, going from very minor severity, such as we would assign to a flash of unjustified indignation, all the way to the infinite severity that many claim attends cursing God. Consider also the set, P, of all possible degrees of punishment, ranging from very mild discomfort of short longevity to the everlasting torments involved in the traditional conception of Hell. We saw that a degree of punishment could be modelled as a set of pairs of intensity (of discomfort, or whatever) and duration. And that a degree of punishment can count as infinite in virtue of the durations it contains. Finally, consider the retributive function,  $f_r$ , that matches up strengths of severity to degrees of punishment. This is the function we all work with when we talk about what punishments deserve. S will therefore be the domain of this function. We can now render (FF) yet more perspicuously:

(FF'') For any collection of moral failings that is both (i) finite in quantity and (ii) finite in severity,  $f_r$  will not assign to that collection an infinite degree of punishment.

### 6.7.2 Pan the Hedonistic Immortal

I will give the following argument via thought experiment to show that (FF'') is false. I will argue that there are cases where an infinite punishment is clearly inadequate to handle an offence of finite severity. But if certain infinite punishments are clearly not strong enough to be adequate punishments, then we have no reason to think that no infinite punishment can be adequate

punishment. The opposition to infinite punishments comes from the conviction that any infinite punishment would be overkill for any offence of finite severity. If that claim is false, then the objection to infinite punishments has little to hold it up.

Consider, then, the following thought experiment. Pan is an immortal god who lives a life of unremitting hedonism and pleasure-seeking amongst the stars. One day he encounters Planet Earth. He is initially tickled by the antics of the planet's inhabitants, but he quickly grows tired of them, and eventually sets the whole world ablaze, guaranteeing a painful, fiery demise for every inhabitant. This is an act of great wickedness, but, for all its wickedness, it surely remains an act of finite evil, because the inhabitants of the Earth are of a finite number.

One other fact about Pan should be noted. Because of his hedonistic nature, he possesses a harem of one hundred naiads, and they accompany him wherever he goes. Furthermore, he has a natural right to this harem, similar to the way that a father has a natural right to his son's obedience.

Suppose Pan comes before God for judgment for his crime. But all God decrees by way of punishment is to take away from Pan one of his naiads. Naiad no. 65 will be permanently, and thus for ever, removed from Pan's entourage. Pan therefore loses his natural right to this naiad, and that ray of elfin sweetness she shone into his life has is forever gone. This is clearly an infinite punishment, for the duration in the punishment pair is infinite:

<loss of access, and right to, naiad no. 65; for ever>

For the rest of Pan's immortal existence, he will be without that member of his harem.

But the problem is that this appears plainly inadequate as a punishment for Pan's misdeed. It just wouldn't, we can imagine, affect him severely enough to be a sufficiently serious punishment. Certainly, he has lost one member of his harem, but he has ninety-nine left to him, so it is no great loss, and his indulgent misadventures in time and space can continue as before.

Not merely are some infinite punishments inadequate, but also some finite punishments are plausibly too strong. Suppose that God sentenced Pan to be tormented in fire for a trillion years instead. There are (rounding up) 8 billion people on Earth at present. Even if the painful fiery demise that Pan caused them to suffer took an hour to culminate in death, that would only amount to 8 billion hours of suffering caused by Pan. To cause Pan to suffer for a trillion *years*, therefore, in punishment, is plausibly an excessive punishment. Not by virtue of the intensity of the pain—we can suppose Pan only suffers the qualitatively same pain he caused the inhabitants of Earth to experience—but by virtue of its duration. *a fortiori*, an infinite punishment of that sort would likewise be too much.

But if it is unjust for Pan to suffer for a trillion years in this way, then it is likewise unjust for him to suffer in that way forever. So, if we know both that (i) certain infinite punishments are not serious enough (because of weak intensity) and that (ii) certain infinite punishments are excessive (because that intensity (given that duration) is too great), then surely it is simply a matter of finding the sweet-spot between these two alternatives: of finding the level of intensity of punishment of infinite duration that is neither too strong nor too weak. It would be strange indeed if there were an abrupt jump from a punishment that were too weak to be adequate punishment, even though of infinite duration, to one slightly stronger that became impermissible to be dispensed over an indefinite duration. Therefore, (FF'') is probably false:  $f_r$  will assign to Pan's offence a degree of punishment that contains at least one pair that contains an infinite duration.

### *6.7.3 Partial and Fully Infinite Degrees of Punishment*

One helpful distinction one should note at this juncture is the distinction between partially infinite and fully infinite degrees of punishment. A degree of punishment is fully infinite iff all the punishment pairs it contains are of infinite duration. A degree of punishment is partially infinite iff it contains at least one pair that contains an infinite duration. In Pan's case, it appears that  $f_r$  would map the severity of Pan's offence to a partially infinite degree of punishment. I argued that there would be an acceptable infinite punishment pair, but it is also surely true that there would be an adequate finite punishment pair, in which case the degree of punishment the offence merits will be partially infinite.

A lot of Christians who believe in Hell (traditionally understood) would claim that every wrongdoing is a wrongdoing against God, and would merit, on that account, eternal suffering. Every wrongdoing would therefore, on their view, merit a fully infinite degree of punishment, no finite punishment ever being sufficient.

### *6.7.4 Some Objections*

One might object to the argument from the Pan case as follows: Pan isn't really punished at all. For, in order to be punished for something, you must experience some discomfort. But Pan doesn't experience any discomfort. He merely endures a slight reduction, a slender diminution, of his daily delight. Bawulski contends for this line of thought. He appears to insist that retributive punishment

requires unpleasant experience. He therefore says that ‘It is hard to see how we might punish an offender who is in a coma, especially if that coma were irreversible. We might be able to extract compensation from her estate, but we would normally consider this means of punishment to be a contingency-plan sentence in lieu of a punishment that involved the offender’s knowledge and recognition of her wrong-doing.’ (2010: 66). And he consequently moves that annihilation, being put out of existence, cannot function as a punishment: ‘there is great difficulty in seeing how annihilation *per se* is punishment at all.’ (2010: 66).

I disagree with Bawulski here. It seems to me quite plain that annihilation can function as a punishment. Suppose an angel indulges in malicious thoughts against another angel. Then the offending angel goes to sleep. God decides that he will punish the angel, and his punishment is that, before the angel wakes up, he puts the angel out of existence—annihilation. The angel is therefore at no point aware of God’s displeasure, nor at any point experiences anything unpleasant. It simply never wakes up. It is forever gone. I find it evident that such an annihilation is more than adequate to function as a possible punishment. Indeed, some might say that, in the angelic case here described, the punishment is too excessive, while Bawulski’s complaint appeared to be that non-existence could never have enough clout to be a punishment.

So, what things can count as punishments, if an unpleasant sensation, or even awareness that one is being punished, is not necessary? Here follows a list of possible ways to punish.

(i) Reduction in happiness. Like the Pan case, this need not be attended with suffering; one may not miss the pleasures one might have otherwise had, but even if there is suffering together with the reduction, the reduction is still part of the punishment in its own right, not merely by virtue of the unhappiness it may cause.

(ii) Loss of a right one formerly possessed. One might lose the right to free speech, say. No suffering need be involved, yet the loss of a right remains a bad thing to happen to one.

(iii) Annihilation. Being pushed out of existence is a serious thing to happen to one, even if the process is painless, and one isn’t feeling anything at all.

(iv) The loss of a good that one would otherwise have had. I don’t think one needs to have a right to a good in order for one’s being deprived of it to function as a punishment. Again, I don’t even think one needs to know that one has been deprived of it for it to be a punishment. Suppose that, as things are proceeding, I am on course to receive a large inheritance some years down the line, though at present I have no right to it. I know nothing of this, however. But, on account of various sins, God deigns to providentially alter the course of history so that I will no longer inherit this fortune. Again, I know nothing of this. But it still seems to me that were I deprived in such a

manner of a fortune that were coming my way, then that would be a punishment. Not a punishment of a very great sort, you say. Perhaps not, but a punishment nevertheless, I feel.

It is also worth noting that being punished can even increase your net happiness. One of the best ways of seeing this is to consider the punishment of being reduced to a state of mental retardation. To be made intellectually disabled is to lose a great many goods, but rather than making you sad, it might make you happier. A chuckling simpleton may lead a much happier life than a tortured genius; nevertheless, to reduce the latter to the former would be a sore punishment. The thought here is that the good one is deprived of is so great that one still suffers a net loss in goods, even though one's happiness might increase. After all, happiness is only one sort of good.

(v) The introduction of an evil into one's life. This is very general, and it includes pain, because pain, I assume, is an evil. Illness, death, destruction, sorrow, ugliness, etc. are all also included under this category. In fact, I say, between (iv) and (v) we cover every possible punishment. I think that anything that functions as a punishment has to involve either the loss of a good or the introduction of a bad into one's life. Annihilation would be the loss of future goods, fiery torment the introduction of a bad, and so on.

All this goes to show two things: first, that Pan's punishment is a genuine instance of punishment (even if not adequate punishment); and, second, that punishment can't simply be a function of duration and intensity of *suffering*—there are more ways to be punished than just through suffering.

### 6.7.5 *The Significance of the Falsity of FF*

Here is a fresh, more perspicuous, statement of (FF):

(FF'') For any collection of moral failings that is both (i) finite in quantity and (ii) finite in severity,  $f_r$  will not assign to that collection an infinite degree of punishment.

We argued for its falsity using the Pan case as follows: if a collection of moral failings satisfying (i) and (ii) can never merit an infinite degree of punishment, then it will be because any infinite degree of punishment would be excessive. But there are clearly infinite degrees of punishment that are non-excessive, indeed, insufficient, to deal with the collection containing only Pan's moral failure. But if some infinite degrees are not excessive, then (a) there is no big worry that pushes us toward accepting (FF''), and (b) if there are non-excessive infinite degrees, then all one needs to do, one

feels, is to increase the goods lost or evils introduced bit by bit until one reaches an infinite degree that is neither insufficient nor excessive.

Indeed, on the strength of this reasoning, I am inclined to think, not merely that (FF") is false, but that (Anti-FF) is true:

(Anti-FF) For any collection of moral failings that is both (i) finite in quantity and (ii) finite in severity,  $f_r$  will assign to that collection a partially infinite degree of punishment.

It seems to me that any offence, no matter how trivial, can be justly punished with an everlasting punishment just so long as you make the loss of goods comprising that punishment smaller still, and therefore there will be at least one punishment pair in the degree that has infinite duration.

With that, the arithmetical certainty with which this principle was tossed about—how could finite sins merit infinite punishment?—is turned on its head. Every finite sin merits a (partially) infinite degree of punishment!

Although this is a rhetorical victory, the typical defender of the justice of Hell has work remaining to him. 'Hell' was defined using (H1–H4), and, as I noted, such a conception is compatible with Hell being not such a bad place, and indeed even a very pleasant place. In order to get closer to the traditional conception, we needed (H5), the Torment Thesis—the contention that to be in Hell is to suffer greatly. That gave us 'Hell+'.

I think the Pan counterexample and what follows from it succeed in showing that there is no obstacle to supposing Hell is just, but that is too much of a low bar. It is the justice of Hell+ that people want to see exonerated. One opposed to the justice of Hell+ could modify the argument given above against the justice of Hell. He could amend the argument given above, as follows:

(1+) Hell+ is a perfect expression of justice: you suffer neither more nor less than you deserve in Hell.

(2+) You are punished only for the sins of your earthly lifetime in Hell+.

(PP) For any culpable fault  $F$  exemplified by an agent  $A$ , there is a *degree* of punishment  $D$  that  $A$  deserves, and  $D$  is proportional to the *severity* of  $F$ .

(3) The sins of our earthly lifetime are all finite in severity.

(4) The sins of our earthly lifetime are finite in quantity.

(FF2) Sins that are of both finite severity and finite quantity cannot merit a degree of punishment that is of infinite duration *and great intensity of suffering*. (From PP)

(5+) Hell+ is a punishment of infinite duration *and great intensity of suffering*.

Therefore,

(6+) Hell+ is not just, *contra* (1+). (From 2+, 3, 4, FF2, 5+)

This argument is not as good as the last one insofar as it is now conceded that (FF) is false, for now the move from (PP) to (FF2) lacks the appearance of arithmetical certainty that gave the move from (PP) to (FF) its original force. It appeared a matter of simple mathematics that offences finite in quantity and severity could not be proportional to an infinite punishment. With that conviction stripped away, we realise that the move from (PP) to (FF2) rests entirely on the inclusion of ‘great intensity of suffering’.

How might the move now be justified? I think it must rest on bare moral intuition. I do not think this is unreasonable. No matter how great and how numerous a man’s earthly crimes, if they remain finite in number and severity, it would be unjust for him to be painfully tormented for them forever—intensities above a certain line justice cannot join with an infinite duration. This would be analogous, perhaps, to the way in which certain horrors, some might say, can’t be permissibly brought about even if outweighed by far greater goods. There is a cap on these matters, and we can only go so far, permissibly.

One way of attacking (FF2) would be to consider cases where the finite offences were gradually raised in quantity and severity until they were so great that eternal suffering seemed an apt punishment, but I won’t pursue such a line because it would probably lead to clashing intuitions, and, more fundamentally, the focus of the objector to the justice of Hell+ isn’t on possible sins of incredible, though finite, severity, it is on the sins of our earthly lives, which don’t strike him as of a collected severity severe enough to warrant everlasting torment.

It is to the undermining of that impression that I now turn. I shall do it by motivating the denial of (3), and then of (4).

## **6.8 Against Our Sins Being Finite in Severity**

There is an established tradition in Christian thought that denies premise (3) from the above argument against the justice of Hell. The tradition says this: not all our sins are of finite severity, because at least some of our sins are against God, and to offend against God is to commit an offence of infinite severity, because he is a being of infinite greatness. In the literature, this idea is referred to as the ‘Status Principle’—the idea that the severity of your wrongdoing is partly a function of the status of the offended party.

### 6.8.1 Anselm on the Status Principle

Here is what Anselm has to say about how serious a matter it is to offend God. In his *Cur Deus Homo* he asks how serious a matter it would be to turn one's head to look in a particular direction when God has forbidden one so to look. The respondent in *Cur Deus Homo's* dialogue says this:

When I consider the action in itself, I see that it is a very slight one; but when I enter fully into what it is when done against the will of God, I see that it is something very serious, and above comparison with any loss whatsoever (1890: 50)

This conclusion is affirmed in the protagonist's voice. He agrees that

Thus gravely do we sin every time we knowingly do anything, however small, against the will of God (1890: 51)

And it follows from that that no action of ours

could suffice to make satisfaction for one sin, however small, when that one act is considered as opposed to the will of God. (1890: 50)

The language of infinity is not used by Anselm, but introducing it would help join the dots. Why can we make no satisfaction for one small sin? What prevents it? Because all sin is an offence against an infinite being, it therefore requires an infinite restitution, something finite human beings are unable to provide.

### 6.8.2 Jonathan Edwards on the Status Principle

Jonathan Edwards is more explicit in introducing both infinity and Hell into the discussion. He writes, in a discourse entitled 'The Justice of God in the Damnation of Sinners',

God is a being *infinitely* lovely, because he hath infinite excellency and beauty. ... He is a being of infinite greatness, majesty, and glory; and therefore he is infinitely honourable. ... His authority over us is infinite; and the ground of his right to our obedience is infinitely strong ...  
So that sin against God, being a violation of infinite obligations, must be a crime infinitely heinous, and so deserving infinite punishment. (1974a [1834]: 669)

After all, Edwards continues,

Nothing is more agreeable to the common sense of mankind, than that sins committed against any one, must be proportionably heinous to the dignity of the being offended and abused (1974a [1834]: 669)

It therefore follows that

If there be *any* evil of faultiness in sin against God, there is certainly *infinite* evil: for if it be any fault at all, it has an infinite aggravation, *viz.* that it is against an infinite object. If it be ever so small upon other accounts, yet if it be any thing, it has one infinite dimension; and so is an infinite evil. (1974a [1834]: 669)

Hell therefore involves no violation of justice:

The *eternity* of the punishment of ungodly men renders it infinite: and it renders it no more than infinite; and therefore renders no more than proportionable to the heinousness of what they are guilty of. (1974a [1834]: 669)

### 6.8.3 Analysing the Status Principle

So much for the sources. But what are to make of the principle? Suppose we state the principle as follows:

(SP) The severity of one's wrongdoing is partly determined by the status or importance of the party offended or wronged by one's wrongdoing (if there is such a party) such that the greater the status of the offended party, the greater the severity of one's wrongdoing, *ceteris paribus*.

Some have objected to the principle on the ground that it relies on a feudal perspective of matters.<sup>41</sup> For a serf to strike his lord was a far greater crime than for him to strike his fellow serf—his equal rather than his superior. But, it may be urged, we have moved on from such an inegalitarian understanding of social relations.

It might be responded that our moving on, for all its progressive, enlightened allure, was not a movement in the right direction. But there is a more decisive response. We all seem to grant that, as far as greatness of being goes, the status principle holds. Consider: to kill a fly because it irritates you is permissible. To kill a dog because it is irritating you, however, looks reprehensible. Finally, to kill another human being because he is irritating you is universally regarded as a terrible thing to do. What explains these different reactions? Difference in greatness of being. A fly is not a being possessed of any great magnitude of greatness or intrinsic value. And that is why to kill it is no great evil. But a dog is a far greater being than a fly, and for that reason we realise that stronger reasons for killing one are necessary than in the case of the fly. Human beings, of course, being made in the image of God, possess a value, and a worth, and a greatness that far outstrips both dogs and flies. Accordingly, reasons of a very great sort are required to be justified in killing another human being.<sup>42</sup>

But if greatness of being is indeed morally relevant to the severity of wrong, then it looks as if the defender of Hell+ has an important principle to appeal to in their defence of the justice of Hell+. It is as Anselm and Edwards have it: God is a being infinitely great, and so our sins against him, being proportional, in part, to the greatness of being we offend, are plausibly infinitely severe. And if they are infinitely severe, there can surely be no obstacle to supposing them to require infinite punishment of a *severe* sort, the sort that Hell+ requires.

I want to discuss some remarks that Kvanvig makes, because they will help clarify matters. He writes,

---

41 Marilyn McCord Adams declares the status principle ‘highly implausible’, and suggests that Anselm’s contrary opinion can be explained simply because he was a creature of his time: ‘it is understandable how Anselm could have come to hold such a principle. Anselm was a member of feudal society in which the amount of honour due to serfs as opposed to lesser nobles, and to lesser nobles as opposed to the king, was very important in dictating behaviour.’ (1975: 442). But, of course, we are all creatures of our time. Kvanvig makes a similar remark to Adams: ‘Some easily grant [the status principle] because their moral experiences involve participation in inegalitarian societies. [...] That these experiences are grounded in the moral dimension of life rather than in the unprincipled and transient character of societal organization is far from obvious.’ (1993: 29). But suppose that is correct. By the same token, it would appear far from obvious that these experiences are grounded in transient societal organisation rather than moral reality.

42 Kvanvig concedes, along similar lines, that it is possible to motivate the principle this way, apart from appeal to societies with structured classes: ‘One such example [showing this] is that the moral guilt, if any, incurred by killing a plant is quite different from that incurred by killing a human being.’ (1993: 30).

even the smallest amount of harm might deserve an infinite punishment. How could this be? According to the status principle, the punishment deserved is determined by two factors; the status of the being affected and the amount of actual or intended harm involved. (1993: 30)

He elaborates on this:

Which sin or how much sin must a person commit against God to justify infinite punishment? There is a plausible argument that one sin alone done against God must be enough if the equal punishment version of the strong view is to be defended. [...] If one sin alone is not enough to warrant infinite punishment, no function on number of sins will do either, for any such function would be completely arbitrary in drawing the line where it did. For example, suppose the claim is that sixteen sins are not enough to warrant an infinite punishment, but seventeen are. The natural response is to wonder why that would be so. We would be at a loss to find any principled reason for drawing the line at a particular number of sins. (1993: 31)<sup>43</sup>

Kvanvig in these passages considers whether the status principle commits one to the view that one sin alone against God is enough to ‘deserve’ or ‘justify’ infinite punishment. If we take ‘deserve’ to mean ‘demand’, such that the claim is that one sin against God cannot be justly punished save by everlasting punishment, then the claim can be stated as follows:

(7) For any sin against God,  $f_r$  will assign to the severity of that act a fully infinite degree of punishment.

If we take the claim that one sin against God justifies infinite punishment to mean that infinite punishment is permissible, though perhaps not required, then we can state this claim as:

(8) For any sin against God,  $f_r$  will assign to the severity of that act a partially infinite degree of punishment.

---

43 What Kvanvig mentions there as the ‘the equal punishment version of the strong view’ is the contention that every occupant of Hell is tormented for ever and to an equal degree. It therefore presupposes the idea that every human being is just as guilty as any other. I take this view to be something of a hyper-Protestant curiosity. The Protestant aversion to the Roman-Catholic belief in justification through meritorious works, and the sort of boasting that is believed to result—‘O wastrel and profligate, my works are greater than yours!’—is carried over to Hell as well. No denizen of Hell will be able to boast that he is less wicked than his peer—they must all be equals in iniquity. This view must be rejected because it is at odds with Scripture. At the Lord’s coming, ‘that servant who knew his master’s will but did not get ready or act according to his will, will receive a severe beating. But the one who did not know, and did what deserved a beating, will receive a light beating.’ (Luke 12:47–48). See also Matt. 11:22 & 24.

(7) entails (8), given the definitions given above, though the converse is false. The defender of Hell+ doesn't strictly need the truth of (7), though it might be supposed he does at first blush. Although if an eternity of torment is held to be just as bad as some finite degree of punishment, then that finite torment must be sore indeed. Indeed, suspicion about the adequacy of any such finite torment is perhaps what drives the defender of Hell+ to prefer (7) over merely (8). It seems that the greater the torment in an infinite punishment pair that can be justly applied, the more sceptical we are about the possibility of any finite punishment pair being just as bad (a member of the same degree of punishment).

But implied in Kvanvig's remarks is perhaps an argument that the status principle issues in absurdity. Consider a holy angel, called 'Phanuel'. Phanuel is one of God's trusted servants, and through the long years of his angelic life he has served God unwaveringly and with perfect obedience and joy. In short, he has been utterly without sin. But, one day, a burst of resentment against God enters his mind. Perhaps that burst of resentment is involuntary, and so Phanuel is not guilty for it. But let us further suppose that, for a second or so, Phanuel consciously decides to indulge this resentment against his maker. Then he comes to, and banishes all such feelings from his mind. His angelic life then proceeds as before, entirely holy and spotless. The status principle, because of the arguments Kvanvig gave, appears committed to the claim that this slender sin on Phanuel's part merits eternal hellfire. But isn't that counterintuitive?

#### 6.8.4 *The Model in Greater Detail*

In order to deal with arguments of that sort, we must develop the model given so far in greater detail. In particular, it should be noted that the status principle is a principle governing *severity*, not one that governs punishment, *per se*.

We introduced  $f_r$  above as a function that takes us from a measure of severity to a degree of punishment. But the status principle governs the function taking us from wrong action types to levels of severity. We can call this function  $f_s$ . We therefore have two functions and three sets. We have  $W$ , the set of all wrong action types;  $S$ , the set of all levels of severity of wrongdoing; and  $P$ , the set of all degrees of punishment. The domain of  $f_s$  is  $W$ , and its range is  $S$ . And, as mentioned above, the domain of  $f_r$  is  $S$ , and its range,  $P$ .

It should be made clear just how  $S$  should be distinguished from  $W$ . The wrong act-types of  $W$  are individuated sufficiently finely that everything morally relevant to the status of the act-type is included minus the moral severity, and therefore the moral wrongness itself, that pertains to the act.

An act-type such as ‘breaking a promise’ is insufficiently fine-grained because much will depend on who the promise was made to, how sincerely it was made, what it concerned, how much faith the recipient of the promise is known to put in it, and so on. All that detail must be included in the act-type, if it is morally relevant.  $W$  will contain all and only those wrong action-types that are maximally morally sensitive in that sense. Severities, on the other hand, appear to be somewhat bare moral properties. How does the severity of unjustifiedly insulting a stranger differ from that of unjustifiedly killing a stranger? No informative answer can be given, it appears, aside from a difference in intrinsic strength. Yet one shouldn’t forget the important role that severity plays in motivating action. Why do I seek satisfaction against the thug who slapped my grandmother? Not merely because he slapped my grandmother, but because slapping my grandmother is a very bad thing to do—it is morally severe to a significant degree.

How do we determine what level of severity any given wrongdoing should be matched up to? What informative rule, in short, does  $f_s$  embody? I’m not sure much can be given in answer to this question. We simply recognise that different levels of severity should be assigned in response to different offences. We just ‘see’ that it is more morally severe to kill someone than to insult them, and so on. We must rely on our moral intuitions to guide us here.

Yet some informative principles of a general sort can be laid down. It might be that a defender of Hell+ might try to exploit the absence of much informative here and take a brutist line of thought. They might say, ‘ $f_s$  assigns to sins against God an infinite severity. Why? It’s just a brute fact. We can no more say why it is that sins against God are infinitely severe than we can say why it is that (unjustifiedly) beating one’s wife has the severity that it has. These are just brute matters.’

But there is at least one principle that undermines, and makes unnecessary, the appeal to brute fact. Call it the Infinite Severity requires Infinite Aspect principle:

ISIA:  $f_s$  assigns to any given morally wrong act-type,  $w_x$ , an infinite severity if and only if  $w_x$  is infinite in some (bad) morally relevant respect.

How does this work? Well, if I push someone into Hell because they annoyed me then that act-type has reference to something infinite: Hell. Or if I cause an infinite number of people pain, then the sum total of the pain I have caused will be infinite. These count as evidently morally relevant infinite aspects of the wrong action. Just about every action, however, will be infinite in some respect. Suppose I bruise someone’s face. The bruise I bring about has an area of 3 cm<sup>2</sup>, let’s say. But if I bring about a bruise of 3 cm<sup>2</sup>, then I have also brought about a bruise of 2.5 cm<sup>2</sup>, and a bruise of 2.25 cm<sup>2</sup>, and of 2.125 cm<sup>2</sup>, and so on *ad infinitum*. I have brought about a bruise with

infinitely many proper parts, in other words. But such an infinity, though present, is also obviously morally irrelevant. I don't have to answer separately for bringing about all those different shaped bruises, just for bringing about a bruise of 3 cm<sup>2</sup>. Again, I can't think of a particularly informative way to demarcate morally relevant infinite aspects of action from ones that are not morally relevant—we can do little more than let our intuitions be our guide.

We can see (ISIA) as following from a more fundamental principle, the Severity Proportionality Principle, which holds that the severity of an offence is proportional to the amount or number of morally relevant respects in the wrong act:

SPP:  $f_s$  assigns to any given morally wrong act-type,  $w_x$ , a severity proportional to the amount or number of (bad) morally relevant respects pertaining to  $w_x$ .

It plausibly follows from SPP that ISIA is true. For suppose that an offence with infinitely many (bad) morally relevant respects was assigned a finite severity,  $s_x$ . Surely it would be possible to reach any possible finite severity by increasing the number of morally relevant bad respects of one's wrongdoing. Deciding (unjustifiedly) to burn one person alive is bad. Deciding (unjustifiedly) to burn two people alive is twice as bad. We can proceed in this manner indefinitely, and thereby increase the severity of the wrongdoing indefinitely. At some point, therefore, there will be a wrongdoing,  $w_x$ , with a finite amount of morally relevant bad aspects that is assigned  $s_x$  by  $f_s$ . But then when we consider a wrongdoing with a greater number of morally relevant bad aspects (which there surely is, for the quantity in  $w_x$  is finite),  $w_{x+1}$ , say, it would be assigned a severity greater than  $s_x$ :  $s_{x+1}$ . But if a finite wrongdoing is assigned a severity greater than an infinite wrongdoing, then  $f_s$  is not assigning severities proportionally, *contra* (SPP).

### 6.8.5 The Case of Phanuel

What then is to be said of the case of Phanuel, the otherwise spotless angel who indulged, for a second or so, a feeling of resentment against the Lord?

The best response that can be made to this, I think, is to note that not all infinite severities are equal. Cantor's work was the breakthrough work that showed that infinities come in many sizes. But Cantorian cardinality would not be helpful for modelling differences in infinite severities, because in order to get one infinite severity to be greater than another infinite severity, Cantorian cardinality requires it to be infinitely greater. It is better therefore to use an ordinal measure.  $\omega$  is the

smallest infinite ordinal—the order type of the natural numbers—and an ordinal measure of size tells us that  $\omega < (\omega + 1) < (\omega + 2) < (\omega + 3) \dots$  etc. Ordinality sees differences when cardinality sees none. In this case, it sees finite differences between infinite quantities.

Mapping infinite severities onto infinite ordinals therefore allows us to see greater and lesser infinite offences. If the status principle is right, then one way of performing an infinitely severe offence is to offend against God. But that alone doesn't tell you what infinite ordinal is assigned to your offence. If your offence were one of brief indifference to God, then that would be a low infinite ordinal,  $\omega + 2$ , say. But if the offence were one of shrieking and persistent blasphemy, then that would be assigned a much higher infinite ordinal,  $\omega + 200$ , perhaps. Differences in infinite severities need to be acknowledged in any case. Intuitively, pricking an infinite number of people in the arm is not as bad as dismembering an infinite number of people. Infinite ordinals look as though they present a good way of capturing these differences in infinitely severe offences.

Now we can address the case of Phaniel, who indulged resentment for a brief while against God. I think that the defender of the status principle should say that Phaniel has committed an infinitely severe offence. I find Kvanvig's argument carries force on this point. For suppose that Phaniel's offence was finitely severe. What, then, would one have to do to commit an offence of infinite severity against God? Presumably it would have to be a more persistent and grievous offence than brief resentment. But, as Kvanvig argues, any line that is drawn between finitely severe sins against God and infinitely severe ones would then appear arbitrary.

I think that the charge of arbitrariness can be rebutted. Consider that if you exert a modest amount of effort throwing a rock, then it will go so far. Perhaps 3 metres. But if you really heave the rock, maybe you can get it to travel 20 metres. But no matter how hard you throw the rock, it will come back to the earth some finite distance away. You might generalise on this account and say that no matter how hard the rock is thrown, it will come back to the earth some finite distance away. But this is false. If it is thrown with sufficient force it will be loosed from the earth's gravitational pull and travel throughout space for ever. This isn't arbitrary because the fact that a certain finite threshold of force, over which results in an infinite distance, is explained in terms of more fundamental laws of force and motion. A defender of the status principle might say a similar thing: there is a certain threshold which, when one crosses it in one's offences against God, they start to count as infinite, but this isn't arbitrary because that fact is explicable in terms of more fundamental moral laws and principles.

However, there is a cousin of Kvanvig's argument that packs more of a punch. Instead of focusing on arbitrariness, one might focus on a 'too small to make a difference' principle. Even if we allow the suggestion that there is a threshold over which offences against God start to count as

infinitely severe (on the basis of more fundamental moral laws), one can still protest against the idea on the ground that it is intuitively wrong for such a small difference (the difference between reaching the threshold and coming just shy of reaching it) to have such strikingly different severities, viz. infinite in the former case and finite in the latter. We can ‘see’, via moral intuition, that morality doesn’t work that way. This is a better way to press home Kvanvig’s thought. Intuitively, therefore, every offence against God must count as infinitely severe.<sup>44</sup>

Not merely that, I also think that Phaniel’s offence *demand*s a punishment of an infinite degree. I think, per (7), that the status principle must have the consequence that every offence against God requires a fully infinite degree of punishment. For suppose that Phaniel’s offence could be paid for by a finite degree of punishment—a few hours of pain. It is perfectly possible to merit a degree of punishment that bad by a misdemeanour of finite severity. But it is very peculiar that an infinitely severe offence should be paid for by a degree of punishment that a finite offence is also paid for by. Indeed, one can even earn a *stronger* degree of punishment than that meted out to Phaniel by increasing the severity of a finitely severe offence until it earns a slightly more severe punishment than Phaniel suffers. But it is very strange for a finitely severe offence to deserve worse punishment than an infinitely severe one.

However, if it is insisted that every infinitely severe offence *demand*s an infinite degree of punishment, then that mollifies the objection to some extent. But it doesn’t entirely remove it. For even though it may remain a peculiar feature of infinite severities that they are only assigned fully infinite degrees of punishment, it will still be the case that some of those punishments will be quite weak. To suppose that Phaniel deserves an eternity of hellfire, say, for his crime seems to us excessive, even though his offence strikes against God, an infinitely great being. But the status principle can accommodate that. Phaniel must simply deserve an infinite punishment that is less severe than that, such as an eternal itch, or annihilation, or an eternal diminution in happiness.

This is good insofar as it goes, but the one remaining problem is that if the infinite punishment Phaniel deserves is slender enough, then it will be possible for one to merit it through offences of finite severity, as we saw above in the Pan case. Indeed, the infinite punishment that Phaniel merits, if it is simply a diminution in daily pleasure, say, might be far less fearsome a prospect than the degree of punishment some finite sins will merit. And again we have the oddity of an infinitely severe offence being assigned a degree of punishment that is more bearable than some that are assigned to offences of finite severity. This, I think, is the only objection that carries serious clout against the Status Principle.

---

<sup>44</sup> One might push the suggestion that God has so arranged matters that all hell-bound humans are clearly above the threshold, but this wouldn’t deal with the fact that, were there two wrongdoers, one just shy of the threshold and the other just reaching it, their deeds would receive markedly different severities.

There are other attacks on the principle current in the literature, but I do not rate them highly. They follow below.

### 6.8.6 *Attacks on the Status Principle: False*

Marilyn McCord Adams objects to the principle because it seems intuitively false to her. She presents the following scenario:

[I]s it true that guilt and liability to punishment are directly proportional, not just to the offence, but to the offended party's worthiness of honour? I think not. Suppose that Schweitzer and Gandhi are equally saintly and that Green and White are equally unsavory characters with long criminal records. Suppose that on separate occasions Green gratuitously slaps Schweitzer in the face, Schweitzer gratuitously slaps White in the face, and Gandhi gratuitously slaps Schweitzer in the face. If guilt were proportional, not just to the offence, but to the moral uprightness of the offended party, then Green would incur more guilt and liability to punishment than would Schweitzer. For since Schweitzer is worthier than White, Green's failure to show respect for Schweitzer was more grievous than Schweitzer's failure to show respect for White. Similarly, Gandhi's action would be more culpable than Schweitzer's. In fact, I think we are more apt to consider guilt as directly proportional to the nature of the offender than to the nature of the offended party. Schweitzer's action in slapping White is, if anything, more culpable than Green's action in slapping Schweitzer. In view of Schweitzer's long-standing habits of self-control and moral behaviour, we should expect more from him than from Green who has never developed these habits. [...] Thus, the principle suggested by Anselm[...]—that guilt and liability to punishment are proportional not merely to the offence but to the majesty of the offended party—seems false. (1975: 443)

But William Wainwright's response here seems decisive:

The principle in question is not clearly false if it is restricted to differences in ontological kinds and not applied to differences between more or less valuable members of the same ontological kind. For consider the following series of actions—destroying a flower, destroying a dog, destroying a human being, and destroying an archangel. Each action in this series appears to be intrinsically worse than its predecessor (presumably because human beings, for example, are a more valuable kind of thing than dogs). But a restricted principle is all we need since God is unique kind of being, and the value of the relevant kind ("divinity") infinitely surpasses the value of the other kinds. (1988: 34–35)<sup>45</sup>

---

45 See also Wainwright 2003.

Thus, one might grant that on an intra-kind level the principle is irrelevant while still providing solid motivation for it at the inter-kind level. But I don't even think that Adams succeeds in showing that the principle fails at the intra-kind level. I think it is true that for Green to slap Gandhi is a worse offence than for Gandhi to slap Green, just so long as we bracket all other facts about their moral characters save for the fact that Gandhi is morally greater than Green. The status principle here motivates a *ceteris paribus* claim: we should think that, *other things being equal*, to offend against a morally greater being is worse than offending against a morally lesser being. But that is just one factor to consider amongst many. As Adams notes, the ease with which the offender's character gives rise to the offence tends to lessen their guilt because they must put up a fight against greater internal proclivities that tend towards committing the offence. Another reason we would be suspicious about granting Gandhi less blame than Green is because of the all-pervasive effects of moral luck. Unsavoury characters such as Green often end up in a life of crime because of a brutal or neglected childhood. Savoury and saint-like characters often live saint-like lives because, being brought up with a silver spoon in their mouths, they have had the luxury to attend to, and develop, their sensitive conscience. Et cetera. The list of all relevant variables might run on to great length, and all Adams's example shows is that the status of the offended party is one variable among many. But no adherent to the status principle should deny that.

Moreover, one can prove that for Green to slap Gandhi is a worse offence than for Gandhi to slap Green, just so long as we bracket all other facts about their moral characters save for the fact that Gandhi is morally greater than Green. For one should, *ceteris paribus*, love what is morally greater more than what is morally lesser. One's love should be responsive to goodness. Therefore, *ceteris paribus*, Green should love Gandhi more than Gandhi should love Green. To act contrary to a duty to love is worse in proportion to the greatness of the demanded love. Therefore, Green would be acting contrary to the greater demanded love.

#### 6.8.7 Attacks on the Status Principle: Incoherent

Seymour (1998) protests that the very notion of an infinitely great being is incoherent. He quotes approvingly Cleanthes from Hume's *Dialogues on Natural Religion*:

I have been apt to suspect the frequent repetition of the word infinite, which we meet with in all theological writers, to savour more of panegyric than of philosophy; and that any purposes of

reasoning, and even of religion, would be better served, were we to rest contented with more accurate and more moderate expressions. (1779: 197)

Seymour considers the notion of an infinitely great being to be akin to the notion of an infinite pain: something of doubtful coherence and possibility. He writes,

[A]ccording to Edwards, it is God's 'infinite greatness, majesty, and glory' which makes him infinitely worthy of honor. Can these traits be infinite? Consider analogous cases. Can something be infinitely beautiful? It might be very beautiful, even as beautiful as something can possibly be, but it is hard to conceive of something being infinitely beautiful. Likewise with pleasure. Perhaps for any pleasure there is one greater that can be conceived; perhaps on the other hand there is a pleasure than which none more pleasurable can be conceived. But what would an infinite pleasure be like? Greatness, majesty and glory are in this respect like beauty and pleasure; they cannot be infinite, not even in God. (1998: 75)

I am inclined to agree with him when it comes to infinite pleasure (and pain), for reasons given previously. But I don't see how the implausibility of an infinite pleasure (which implausibility is derived from consideration of phenomenology) carries over to the supposition of an infinite greatness (which can't be considered suspicious on phenomenological grounds). After all, some things are infinite, such as the number of points in a 1-metre line, and the collection of natural numbers. Why can't God's greatness be counted as one of them?

But we can say more to motivate the infinite greatness of God. It is a given in perfect-being theology that God is the greatest possible being. Whence it follows that God must be a greater being than every finite being. But not merely that, God must be greater than any *possible* finite being. For any possible finite being you can think of—beetles, dogs, cats, apes, human beings, dragons, extra-terrestrials, archangels, Greek gods—God must be greater than them all. But it appears that the set of possible finite beings can be arranged in an order of greatness stretching upward indefinitely. Beetles aren't as great as dogs which aren't as great as human beings which aren't as great as dragons which aren't as great as Greek gods, etc.

Now, the mere fact that God has the property that, for any possible being that is not God, God is greater than that being, isn't sufficient to show that God is infinitely great. For it may be that (i) there is only a finite number of possible beings or (ii) that all possible beings that aren't God have a greatness that is below a certain limit. But reflection on the ordering noted above undermines both these assumptions. Consider that, for any possible dragon, you can surely conceive of a stronger, cleverer, and therefore greater, dragon. This undermines (i). It also undermines (ii), but that requires

a bit more work, because someone might suppose that beings indefinitely increase in greatness, but that that increase approaches a limit, in the same way that the indefinitely increasing series (1.5, 1.75, 1.875, ...) approaches its limit, 2. But the increases in greatness we envision we can perform on possible beings of finite greatness don't appear to get indefinitely 'thinner'; they seem to be as substantial as we wish them to be. Indeed, we can imagine that the increases get greater: consider a dragon with double the strength and intelligence of dragon<sub>1</sub>, and then consider the dragon with double the intelligence and strength of that one, and so on. But God must be greater than all such possible dragons; ergo, God is infinitely great, in the same way that an infinite number is greater than any possible finite number, and Seymour's suspicions are unfounded.

### *6.8.8 Attacks on the Status Principle: Useless*

Kvanvig argues that, even granting the Status Principle, it is a principle of limited use. The defender of the traditional account of Hell wants it to be the case that every unbeliever goes to Hell, and therefore deserves to go there. But if the status principle is how it is explained that people can merit the punishment of Hell, then it needs to be the case that every unbeliever has sinned against God. But Kvanvig considers this a hard sell. He first asks us to distinguish between a restrictive and general view of when a man sins against God:

When does a person sin against God? The two views one might hold here are a restrictive view and a perfectly general view. In the perfectly general view, every sin is a sin against God. In the restrictive view, only some sins are sins against God. If only some particular sins are against God, presumably they would be those in which God is the intentional object of some action. For example, if I throw a rock [~~and~~] to hit my cat with it, then the cat is the intentional object of this action. Just so, the restrictive view about sinning against God would seem likely to claim that one sins against God when and only when one "aims at" or "strikes at" God. (1993: 32)

He then goes on to say,

Defenders of the [...] strong view will find little comfort in the restrictive view of sinning against God, however. The view that every person has at some point explicitly aimed at harming God or struck out at God in some way or other is difficult to sustain. In an age of growing atheism and agnosticism, some people may go through their entire lives never giving God much thought at all.

Yet according to the [...] strong view, all people [are] headed for hell, whether they have given thought to God or not. (1993: 32)

Should the defender of the status principle accept the restrictive view or the general view? I do not find the general view to be very plausible. It is plausible to suppose that one's moral guilt can't extend further than what one is aware of in one's mind. If a child spitefully strikes its sibling, conscious of nothing else but that spite, then I think it implausible to argue that the child is guilty of something infinitely severe, and guilty of something infinitely severe because it has, say, struck a being made in the image of God, and in that way triggered the status principle. No awareness of any of those things was in the child's mind. It seems to me it can therefore justly plead ignorance of any offence against God (though not ignorance of wrongdoing).

I also think it is possible to offend God without triggering the status principle. Consider the following case. Suppose Ivan's job is to monitor packages that are sent through the post. They come by his station at intervals, brought on a conveyor belt. They come through intermittently, so he must be alert to their appearance. If he doesn't check a package as it comes to him, then it might enter the mail incorrectly addressed and be lost in the post. As it happens, unfortunately, Ivan was lazy one day at his post, and a package passed by without him checking it, and it was incorrectly addressed, and therefore lost in post. No one knows where it went, and retrieving it is practically impossible. Ivan is guilty for that, and if his boss decided to take it out of Ivan's wages that seems reasonable enough. But suppose that the owner of the lost package then informed Ivan's boss that the item in the lost package was one of *infinite* value. And suppose this is true. Are we therefore to conclude that Ivan, because of his failure in watchfulness, is guilty to an *infinite* degree? And that his boss is therefore justified in docking Ivan's wages forever? Intuitively not, because Ivan had no idea that the stakes were that high. There was no awareness of any infinite aspect to the matter in Ivan's mind.

Change the case a bit more. Suppose that the lost package didn't contain an object of infinite value, but that the package was directly owned by God, and that Ivan had therefore mislaid God's package. Ivan has therefore wronged God, but again it would be wrong to say that Ivan was infinitely guilty, even though he has wronged an infinitely great being. Because Ivan had no idea that it was God's package—the thought never occurred to him—his ignorance prevents the status principle from being triggered. One needs to have an awareness of God in one's mind in some fashion in order to commit an infinitely grave wrongdoing against God *via* the status principle.

It is such considerations that incline me against the general view. Not every offence, nor even every offence against God, makes one infinitely guilty. Only those done with the relevant sort of conscious awareness.<sup>46</sup> We must therefore prefer the particular view.

How, then, should one respond to Kvanvig's claim that it isn't plausible to think that every non-Christian has at some point committed an offence directed at God in the way that the particular view requires? I see two good responses:

(i) I think one can simply deny Kvanvig's claim that "some people may go through their entire lives never giving God much thought at all" (1993:32). Or rather, one can accept *that* claim. To refuse to give God *much* thought, after giving him *some* thought, is surely a great crime. Once the notion of the deity has crossed one's mind, and one sees its great importance (and how could one not?), and *then* to refuse to dwell on such things anymore is something it is very hard to see how one could do without committing an infinitely grave offence against God *via* the status principle. The more worrying possibility is that the idea of deity should never cross some people's mind at all, so that it isn't possible for them to sin against God in thought. But against that is the long Christian tradition that God has designed man's mind with an inbuilt sense of the deity, the celebrated *sensus divinitatis*. This faculty is what explains the near-universal practices of worship to a creator being we find across the world. Although one might train oneself to suppress the workings of this faculty, it will always have flashed through to one's consciousness at one point or another, inclining one to belief in God and consideration of the divine. But to respond wrongly to such presentations made by the faculty is to wrong God infinitely.

(ii) The second response is to appeal to the doctrine of original sin. One can punt matters back to Eden. The doctrine of original sin (or at least a sufficiently conservative version of the doctrine) has it that every human being is guilty of Adam's trespass. Quite how it can be that such a sin is fairly imputed to us, or how we can plausibly be considered actors of Adam's trespass, is a matter of debate among philosophical theologians. See Rea (2007) for a contemporary take. However, discussion of the plausibility of the doctrine would take us too far afield. Let it just be noted that *if* the doctrine is plausible and the objections made against it can be rebutted, then it doesn't matter if there are people to whom the thought of God literally never occurs. On the Day of Judgment, they won't be guilty of an infinitely severe offence against God on account of what they did between their birth and death, but for what they did in Adam before they were born. And Adam's sin was done in full Edenic splendour. He could not plead forgetfulness or clouded vision. It was an offence

---

<sup>46</sup> This also helps explain why the truth of Psalm 51:4 ('Against you, you only, have I sinned and done what is evil in your sight') is not inconsistent with the claim that some sins we commit are only finitely severe. Even though every sin we commit counts as wronging God in some way, not every way of wronging God is enough to trigger the status principle.

committed with the full knowledge that it was a transgression of the command of an infinitely great God.

### 6.8.9 Summary

For all these reasons, I consider the Status Principle to give a viable understanding of how it can be that Hell (and Hell+) is just. I think its weak points are of the sort brought to light by the case of Phaulon—must even the slenderest sin against God demand infinite punishment?—and of the unevangelised nations—is original sin really doing most of the work in explaining why these people deserve to go to Hell? These are not points of great potency—they might simply be accepted—but it would be better if we had a theory that didn't issue in such consequences. I provide such an account in the next section.

## 6.9 Against Our Sins Being Finite in Number

In this section I will argue that (4) (the claim that 'The sins of our earthly lifetime are finite in quantity') is plausibly false. Or rather, I will not argue that our *sins* are infinite in quantity, but that our *moral faults* are infinite in quantity. If my suggestion is plausible, then it would appear to be a plain resolution of our difficulty. If we are guilty of offending God in an infinite number of ways, then of course God would be just in punishing us infinitely.

### 6.9.1 Blameworthiness for what you would have done

The argument of this section is basically this: you are plausibly blameworthy, not just for the things you actually do, but also for the things that you would do and would have done. But these things are infinite in quantity. Therefore, you are infinitely guilty.

To motivate the idea, consider the following thought experiment. You have a good friend of many years called Alf. You are both dog-owners, and you and Alf are in the habit of exercising your dogs together, going for long morning walks in the country, talking with one another, to that end. On one occasion you are called away on an urgent matter, and can't make your morning appointment with Alf. But, before you leave, you give your dog to Alf, and Alf agrees to give your

dog his morning exercise by himself. As Alf is walking along with your dog, he is approached by a disheveled old man. This old man has a startling proposition for Alf. He offers him £40 pounds if Alf will kill your dog right here and now. Alf refuses, the old man goes on his way, and Alf tells you about the matter when he next sees you. However, the old man was really a jinn, and this jinn also comes to see you. This jinn is very knowledgeable and trustworthy, and you know it. The jinn tells you that although Alf rightly refused his offer of £40, the jinn, on account of his connection to the supernatural, can also tell you this: that if he had offered Alf £50, then Alf would have accepted and your dog would now be dead.

Now, even though Alf has not *actually* done anything wrong, many of us would nevertheless naturally feel a great indignation at Alf were this fact about what he *would have done* to be revealed to us. More than that, we would feel it was a quite *justified* indignation. ‘He would have done *what?! That’s outrageous!*’ Neither is it merely an emotional matter. It would also affect our behaviour towards Alf. We wouldn’t be inclined to let him walk our dogs anymore, for one thing. It might even be sufficient occasion to terminate our friendship with Alf—‘You’d kill my dog for £50, would you? You’re no friend of *mine!*’ You rightly stand in moral judgement over Alf on account of this counterfactual wrong. And justified indignation presupposes the presence of a culpable fault. Therefore, we are responsible, to some extent, for at least some counterfactual wrongs.

### 6.9.2 Counterfactual Wrongs and Characters

One might try to explain the phenomenon in this way: we hold Alf responsible for what he would have done because what he would have done is grounded in his character. And Alf has formed his character into its present state through many years’ worth of decision-making. We do not hold him responsible for the counterfactual wrong, rather we hold him responsible for what that counterfactual wrong implies about him: that he has decided to develop his character in bad ways, or at least that he has refrained from developing it in the good ways that he should have.

But, for one thing, this only pushes the buck a step further back. What if, with those earlier character-forming decisions, Alf would have made terribly worse ones if only he were tempted in a very slight way? Aren’t we still inclined to consider him blameworthy, to some extent, on account on those counterfactual moral failures?

For another thing, I think this problem is really generated by deeper facts about our moral situation than past character-forming. Suppose we are hunting a cruel despot guilty of Hitler-like atrocities. He manages to elude capture for a while, but we eventually trap him. Yet as we stand

over him in triumph, he speaks to us scornfully, ‘How proud and haughty you look! No doubt you fancy yourselves as better, nobler creatures than I. But I can tell you this: that if you were in my position, and brought up in the same circumstances I was, you would have done exactly the same as I did.’ If true, that would rather take the wind out of one’s moral sails. One might think: it was just bad luck for this despot that he was born in those circumstances and I wasn’t. Am I really entitled to claim moral superiority over him?<sup>47</sup> The problem is made keener by considering God’s perspective at the Day of Judgement. A natural expectation to have of God as a perfectly fair and impartial judge is that he would abstract away from such contingencies as were you were born and the nature of your upbringing, and assess your life relative to a broader set of circumstances than merely those circumstances that just so happened to obtain. But then it looks like God would assess you relative to entirely different characters, because a different upbringing is plausibly sufficient for a different character.

But given that God would assess us relative to non-actual and merely possible circumstances and characters, we need to ask how broad the range of possible circumstances is. But it seems the fairest and least arbitrary answer is the maximal one! God will assess you relative to every possible circumstance you might have been placed in and relative to every character you might have been given. But there is an infinite amount of such characters and circumstances. This opens up the possibility, indeed, the expectation, of infinite moral failure.

### 6.9.3 *Counterfactual Wrongs on Molinism*

I shall show how infinite moral failure is an anticipated consequence on both a theological determinist and a Molinist model. I will begin with Molinism. Molinism holds that God, when he was deciding to create the world, consulted his middle knowledge, this vast infinity of subjunctive conditionals of freedom of the form ‘ $S$  would freely  $\phi$  in circumstance  $C$ ’, for any possible agent and any possible circumstance, and created the world (the complete set of persons and circumstances) that gave the collection of free decisions closest to God’s ideal.

Molinism therefore has the implication that, for any existing person, God knows an infinity of facts about what they would do. For any possible life-up-to-a-point and external situation, God knows what any existing person would freely choose to do at that point in time if they were given that life-history and situation. ‘Circumstance’ must be understood in that broad way, to denote a possible-world history. Then the argument proceeds as follows. Consider a world history,  $w_1$  up to a

---

<sup>47</sup> This is the problem of moral luck: doesn’t the fact that what we do and who we are so often depends on factors beyond our control undermine moral judgement? See Nelkin 2019 for an overview.

time  $t$ , that looks more or less like ours up to the present time, although in this world history there is a big difference: you are in a position to commit a great atrocity at  $t$ . Either you decide commit the atrocity in  $w_1$  or you do not. If you do, then you are counterfactually guilty of a great sin. If you do not, then we can adjust the world-history of  $w_1$  just a bit. Consider  $w_2$ , which is just like  $w_1$  save that a hydrogen atom in Alpha Centauri has been moved half an inch to the left. Do you commit the atrocity in that world history? If not, then consider  $w_3$ , in which the hydrogen atom is moved an extra quarter of an inch to the left. If you remain obdurate, consider  $w_4$  in which the atom is moved a further eighth of an inch to the left, etc. etc. Given that this operation can be performed an infinite number of times, the probability of your refraining from performing the atrocity in each of these very slightly different world histories is zero. But that same sequence of argument (moving a molecule about in a distant part of the universe until you get the moral failing) can be run again with any possible atrocity, and given that there are an infinite number possible atrocities to commit, you will be guilty of infinitely many counterfactual moral atrocities—an infinite moral failure.

#### *6.9.4 Counterfactual Wrongs on Theological Determinism*

The conclusion is much easier to reach on a determinist understanding. On theological determinism the sins you commit are determined by the character (or, on my account, by the agential nature (see ch. 4)) that God decides to give you: God doesn't have to give you the particular constitution and circumstance and 'wait and see' if you do will the right thing or not (as on libertarianism); your sins will follow from the nature or character that God gives you. But if we follow the argument of the above sections, then God is going to take into account, when he judges us, what we would do were we given different characters. And, in my terms, what we would do if we were given different agential natures (recall that I affirmed in ch. 4 that one's agential nature is not essential to one). No one can boast, I believe, that a certain agential nature or character is impossible with their personal identity. If that is right, then it will be true of every existing human being that, were they given a certain character or agential nature and put in the right circumstance, they would commit a great atrocity. One can perceive our infinite guilt in two ways, then. First, there is surely a possible agential nature or character that always sins no matter what circumstance it is put in. This perfectly sinful agential nature or character is one we all might have had. But because there are an infinite number of circumstances, there are an infinite number of counterfactual sins attached to this agential nature or character, and it is thereby true of all of us that 'For any circumstance from the infinite set of all circumstances, were we given this agential nature or character, we would sin in

that circumstance’—an infinite moral failure. The other way to see our infinite guilt is to realise that there are an infinite amount of characters or agential natures that will commit a great atrocity in at least one circumstance. It will therefore be true of all of us that were we to be given a particular one of those characters or agential natures and placed in the right circumstance, we would commit a great atrocity. There will be an infinite number of counterfactual truths of that form, each one corresponding to one of those characters or agential natures—again, an infinite moral failure.

Really, the easiest way to make that point that you are infinitely guilty on theological determinism is to note that truths about what you would do, on that understanding, don’t come apart from truths about what it is possible for you to do. What you would do can be determined just by examining the modal space of possibilities. How many possible worlds are there in which you wear a red hat? An infinite number. Likewise: how many possible worlds are there in which you commit a great atrocity? An infinite number.

#### 6.9.5 Objections Regarding the Use of Infinite Counterfactual Wrongs

Thus, we have a fresh way of seeing how the infinite punishment of Hell is just: it is grounded in the infinity of counterfactual wrongs one is guilty of. Now for some objections.

*Objection 1:* You can’t blame someone for something they didn’t actually do. Counterfactual wrongs are precisely that: counter to the facts! They aren’t *real*.

*Response:* We need some way to explain why it is that we are angry at Alf when we find out about what he would have done. But the complaint that we have left reality behind is not justified. The relevant counterfactuals are *actually* true, after all. And it is natural to think that these truths are grounded in some feature of actual reality. For theological determinists, this is easy to see: the truths are grounded in the reality of characters, agential natures, and persons. It is harder to see what grounds the truth of the counterfactuals of freedom on Molinism—and that is one of the famous objections to the system, for the relevant truths about what persons would freely do are true before those persons exist (before creation), and one wonders what *makes* them true.<sup>48</sup> But, at any rate, the Molinist will still hold that all these subjunctive conditionals about our wrongdoing are all actually true, and that seems like a sufficient response to this objection.

*Objection 2:* Very well, I grant that I am guilty of an infinite number of counterfactual wrongs. There are an infinity of worlds where I curse God to his face, and where I torture whole worlds to death for my own amusement. But there another side to the coin. There is also an infinite number of

---

48 See Adams 1977 for a classic statement of the grounding objection to Molinism.

noble and morally superb counterfactual deeds to my credit. There is an infinity of worlds where I endure a million years' torment yet still refuse to murmur against my creator, and a countless number where I selflessly lay down my life so that others may live. Isn't the weight of all my counterfactual wrongs cancelled out by the splendour of all my counterfactual moral magnificence?

*Response:* I accept the counterfactual moral magnificence, but I deny that it cancels out the counterfactual moral transgressions. Here is one reason. It is natural to think that God doesn't expect you merely to fulfil your duties and act well in actual circumstances, but he expects it in any circumstance that might come to pass. But when he sees all the counterfactual wrongs you perform, he realises that you are not meeting his standards, modally speaking. When he sees all your counterfactual good and noble behaviour, he merely sees you as meeting his standards. Thus, in the same way that a lifetime of honest behaviour doesn't prevent one from being tried for fraud, one's counterfactual holiness doesn't outweigh one's counterfactual depravity. Another reason I don't think counterfactual wrongs can be cancelled out in the way suggested is the following. Leave one's actual deeds to the side and consider how God's sees you just in the space of possibilities. He sees that you are just as easily turned to evil as you are to good. For every morally splendid thing you do in some possible circumstance, it is matched by a morally terrible thing you do in some other possible circumstance. You thus exemplify a sort of moral indifference: modally speaking, just as much at home with evil as with good. But surely it is a bad thing to be a person equally at home with evil as with good. This indifference is therefore really an indifference of a morally odious sort—one that contrasts greatly with God's own necessary moral perfection.

*Objection 3:* Doesn't this view imply a sort of futility about the moral life? We can't ever improve our moral standing before God because any actual good deed we do will be matched by a counterfactual wicked deed. All of us are really the same, morally speaking. The infinity of counterfactual possibilities has a great levelling effect, making us all good and evil to the same degree. Moreover, we saw above that God will punish human beings to different degrees. How is this possible on the doctrine you suggest, for it seems to follow that we are all equally wicked?

*Response:* I don't think that a counterfactual murder, say, is as blameworthy an affair as an actual murder. Likewise, I don't think that an actual case of righteousness is weighted as heavily as that same case of righteousness but counterfactual. Even though I think our counterfactual behaviour matters, it does not matter as much as our actual behaviour. I'm sympathetic to a sort of distance principle. I think that the more distant one's counterfactual wrongs are from the actual world—the more fanciful or dissimilar from ours the possible worlds in which they occur—the less they count against us. That, I think, is enough to deal with the problem of moral motivation, and to explain how it is that we not all equally wicked. Of course, one should also remember that it is the

traditional Protestant position that the moral life *is* in a real sense futile: one cannot earn God's favour merely by being moral—the imputed righteousness of Christ is required for that.

## 6.10 Conclusion

We have therefore encountered three ways in which we can see how Hell+ is a just recompense for man's sin. The first was the Continuing-Sin hypothesis—if man is always sinning, then it is little wonder that he is always being punished. That theory is hard to square with Arminian commitments—why are the occupants of Hell always freely choosing to sin?—but not so with Calvinist doctrine, and it is a feather in the cap of the theological determinist that he can explain the justice of Hell+ this way. (Of course, the reason why God would make it the case that some creatures sin forever is a separate question, and one covered in the next chapter.) The second was the Status Principle. Offences against an infinitely great being merit infinite punishment. The principle admitted of plausible defence, though it was slightly weak when it came to handling cases like the Phaniel case—sins of a very minor sort against God—and cases of those who hadn't heard of God. Lastly, we saw that the infinity of counterfactual wrongs that we are guilty of could also explain the justice of Hell+. Again, the theological determinist had an advantage here insofar as an infinity of counterfactual wrongs follows more readily on his account, and also because he doesn't have a worries about grounding these counterfactual wrongs.

That is all I have to say here about the justice of Hell. Even though both Calvinists and Arminians believe in the justice of Hell, they don't both believe that God unconditionally predestined a large portion of humanity of Hell. This raises problems for the Calvinist that are the subject of the next chapter.

## Chapter 7—Calvinism and the Problem of Hell

### 7.1 Introduction

What I intend to do in this chapter is to sketch a theodicy of Hell consistent with Calvinism— that is, with theological determinism and compatibilism about moral responsibility. Many are sceptical of the idea that any such theodicy could be made plausible. We saw in ch. 3 that many critics of Calvinism were baffled by the suggestion that it is God that determines man to sin and for most of mankind (the reprobate) to reject the way of salvation. Present-day writers also make the same complaint. Here are Pereboom’s remarks:

Historically, perhaps the most effective reason for rejecting any sort of divine determinism, and endorsing instead libertarian free will is the unconscionability of God’s damning people to hell after determining them to sin. (2005: 82)

I believe this feeling of unconscionability persists and is widely shared. Consider Baggett and Walls’s complaint against the Calvinist:

[D]amnation involves infinite, eternal misery. For God to choose to consign persons to such a fate when he could have just as easily determined them to joy and happiness [in heaven] is [...] morally obnoxious [...] [It] strikes us as a paradigmatic example of hateful behaviour, not loving behaviour. (2011: 74)

Of course Arminians, being libertarians, deny both that God determines anyone to sin and that it is a straightforward matter for God to ensure people’s entrance into Heaven: the free decisions of humanity, and in particular the decision to accept or reject God’s offer of salvation, upon which entrance to Heaven is conditional, are, at least in some sense, beyond God’s complete control. This belief gives them a great advantage—it permits the deployment of various free-will theodicies of Hell: God greatly desires to save us all from Hell, but it would be wrong for him not to respect our free choice to form, if we wish, a character that is implacably opposed to him, as Swinburne (1983) proposes; or Craig’s (1989) suggestion that although God dearly wishes to save everyone from Hell, when God consults his middle knowledge it so turns out that in all those worlds in which a

significant number freely choose to go to Heaven a significant number go to Hell, and God isn't going to let the fact that many will go to Hell prevent him creating one of those worlds (cf. Seymour 2000).

But no such defence is available to the theological determinist, for on his view God has complete control over what human beings will choose, because he is the ultimate cause of all human choices. In short, if Calvinism is true, it seems perfectly easy for God to create a world in which universalism is true—a world in which everyone accepts God's offer of salvation and goes to Heaven. Why wouldn't God do this? What could stop him? 'Surely nothing,' says the Arminian, 'and so Calvinism is false.'

In this chapter I offer what I take to be plausible reasons why God wouldn't, despite a comprehensive control over what his creatures choose, make it the case that all accept the offer of salvation. The reasons I offer are not that novel. They can be found in Jonathan Edwards, John Calvin and across the Reformed tradition in Christian thought. Consider Calvin's remarks on the reprobate:

they were raised up by the just but inscrutable judgment of God, to show forth his glory by their condemnation. (*Institutes*, 3.24.14)

Here again is the *Westminster Confession*, ch. III, 'Of God's Eternal Decree':

VII. The rest of mankind [that is, the non-elect or reprobate] God was pleased, according to the unsearchable counsel of His own will, whereby He extends or withholds mercy, as He pleases, for the glory of His sovereign power over His creatures, to pass by; and to ordain them to dishonour and wrath for their sin, to the praise of His glorious justice. (Schaff, 1977: 610)

Both of those two passages suggest a *motive* for God's decree of reprobation. Calvin suggests that God does it 'to show forth his glory by the reprobate's condemnation', and the *Westminster* suggests that God reprobates for 'the praise of his glorious justice'. We might summarise the thought as follows: important aspects of the divine majesty are not displayed if everyone is saved, so God decrees that many shall refuse his offer of salvation in order that the glories of God's sovereign power and justice might be displayed in their eternal destruction.

But note that displaying is a triadic relation, and we need to ask to whom God is displaying his glories. A traditional assumption is that it is the elect in Heaven, those whom God predestined to salvation. The theodicy I will propose here aligns with this suggestion. I suggest here that it is for

the sake of the occupants of Heaven that God creates people to occupy Hell. It is good to understand God's character and our relation to him, and the occupation of Hell enables both an understanding of God's nature, and good attitudes towards God on the part of the elect that wouldn't be possible otherwise. This chapter will list these benefits.

## 7.2 The Problem of Hell

But let us first state the problem of Hell more carefully. The nature of Hell was discussed in ch. 6, so we don't need to go over that. But what is the *problem* of Hell? As I see it, the problem of Hell is a special case of the problem of evil. Rowe's formulations of the argument from evil have been influential. Let us take the version from his well-known (1979) paper:

1. There exist instances of intense suffering which an omnipotent, omniscient being could have prevented without thereby losing some greater good or permitting some evil equally bad or worse.
2. An omniscient, wholly good being would prevent the occurrence of any intense suffering it could, unless it could not do so without thereby losing some greater good or permitting some evil equally bad or worse.
3. There does not exist an omnipotent, omniscient, wholly good being.<sup>49</sup>

The problem of Hell can be cast in the same mould. Let us put it like this:

4. If Hell exists, then Hell is an instance of intense suffering which an omnipotent, omniscient being could prevent without thereby losing some greater good or permitting some evil equally bad or worse.
5. An omniscient, wholly good being would prevent the occurrence of any intense suffering it could, unless it could not do so without thereby losing some greater good or permitting some evil equally bad or worse.
6. If Hell exists, then there does not exist an omnipotent, omniscient, wholly good being.

The Arminian free-will theodicies mentioned before can be construed as a denial of (4). The greater good God would lose by ensuring there is no Hell is the freely chosen fellowship of his creatures,

---

<sup>49</sup> I am aware of Rowe's later formulations of the argument which Rowe judges to be superior (viz. his 1991 and 1996 pieces). However, I use this presentation of the argument since it better facilitates an exposure of a mistake I take to be commonplace in discussion of the argument from evil: the idea that the justifying good must be greater.

something over which God has no direct control. (Again, the theological determinist cannot make any such move: he believes it is within God's power to bring it about that all creatures freely choose him.) The Calvinist could, however, make some of the standard moves the theist makes in response to the problem of evil: he could appeal to *sceptical theism* and claim that the odds of our knowing God's justifying goods are small, so the fact that we are unable to think of what the goods might be is of little surprise. Wykstra (1984) and Alston (1991) are prominent examples of this strategy. I see no reason for this sort of response to be any less successful here.<sup>50</sup>

But I intend to give a theodicy here. What is a theodicy? Plantinga distinguishes between a *theodicy* and a *defence* (1974: 192): the latter simply suggests broadly logically possible reasons God might have for permitting evil, to show that there is no broadly logical inconsistency between the existence of God and the existence of evil; but to give a theodicy, on Plantinga's rather strict account, is to give the *actual* reasons God has for permitting evil. I find this understanding of 'theodicy' to be too strong. To my mind, to provide *plausibly actual* reasons—reasons that obtain for all we know—is all that the task of theodicy requires. David Lewis agrees with me, saying: "Defense is too easy; knowing God's mind is too hard. I think the topic worth pursuing falls in between [...] [The Christian] can hope to advance from a predicament of not having a clue to a predicament of indecision between several not too-unbelievable hypotheses" (2002: 105–106). And that is the task I set myself here: to offer a plausible explanation of why God would predestine anyone to Hell. I shall adopt the following account of what it is for a reason God has for permitting or bringing about evil to be plausibly actual:

God has a plausibly actual reason R for permitting or bringing about evil =df there is no clearly probative argument against our supposing R to be one of God's reasons for permitting or bringing about evil.

Note that I am not, therefore, concerned to *prove* that the reasons I propose will do the job; rather, I rest content just so long as no-one can prove they cannot do the job.<sup>51</sup>

It must also again be stressed that the problem I am addressing here, the problem of Hell, is very much distinct from the problem that engaged us in the last chapter—the issue of the justice of Hell.

---

50 Though I happen to think the Calvinist should not stop here. I read Paul in Romans 9:22–23 as proposing a theodicy of the same kind I defend in this chapter. Admittedly, Paul's remarks are preceded by a what if', so he who takes the inspiration of Scripture seriously is arguably not committed to the *truth* of Paul's proposal, but surely neither is what Paul says to be ignored as worthless. It seems that the happy *via media* is therefore to take Paul's theodicy as adequate, or true for all we know. So if the Calvinist reads Paul as I do, then he is committed to the *adequacy* of the theodicy I give here, if not its truth.

51 Also, for the record: I myself would not like to be interpreted as claiming that the theodicy I give here is true; my claim is likewise only that I see no clear objection to its truth.

The question about the justice of Hell is this: how could a human being justly merit eternal damnation? And that has been addressed. The problem of Hell is this: how could it be that God has *let* anyone justly merit eternal damnation? Granting that it is morally possible for someone to merit damnation, nevertheless, why has God allowed, or arranged for, something like that to come to pass?

Another way of putting it: the justice of Hell concerns God's pronouncements on the Day of Judgement. How can he justly condemn anyone to eternal suffering? The problem of Hell concerns God's decree before the foundation of the world. Why has God decided to create some people he intends to (or at least knows will) occupy Hell for all eternity? I take it that the moral possibility of a human being justly meriting eternal punishment has been established; I am therefore concerned here with why God would bring it about that any human being did *in fact* merit eternal punishment. I want to list the goods that give God his reasons.

But before I lay out the goods God reprobates for the sake of, let us note that two things are required of these goods. The first is that they must be goods for which the inhabitation of Hell is a necessary condition. The second is to make sure that we can believe they are worthwhile given the evil they require. So I propose that we can offer an explanation of why God would be motivated to reprobate if there is (a) some good state of affairs for which reprobation is a necessary condition, and also that (b) the degree of intensity or amount of the good state of affairs is not so low as to make it obviously not worthwhile given the evil that accompanies it. Only if the goods I propose satisfy both conditions are they goods that would provide God with plausibly actual reasons for reprobation. Also, if a good in question presupposes particular doctrinal or philosophical commitments, then it must also be the case that these doctrinal or philosophical commitments are not too controversial. I shall first suggest the different goods and then discuss the question of their worthwhileness.

### **7.3 Calvinist Theodicy**

I have already given a broad outline of the goods which Calvinists have proposed historically. John Piper is a contemporary theologian who gives a more recent statement of the suggestion:

My answer to the question about what restrains God's will to save all people is this: it is God's supreme commitment to uphold and display the full range of his glory through the sovereign demonstration of all his perfections, including his wrath and mercy, for the enjoyment of his chosen

and believing people [...] This everlasting and ever-increasing joy of God's people in all of God's perfections is the shining forth of God's glory, which was his main aim in creation and redemption. (2000: 339)

Importantly, unlike Piper, I do *not* insist that the displaying of his glories is God's main aim in creation. I don't deny it either, the point is only that the theodicy I give here doesn't require this claim. It requires only the more minimal suggestion that if some state of affairs grants us a clearer or greater or better understanding of the nature of God (or some other beneficial attitude), then God, if he seeks our good, would have reason to actualise that state of affairs (because such a mental state is a good). Whether or not he would grant it because of a more ultimate objective to display the range of his glory is a separate issue. This theodicy does not propose God's glory as the good for which God actualises evil, rather displays of God's glory are here viewed as good because they ennoble or enhance the state of mind of the elect. For the elect to perceive God's glory is good thing in its own right. It is these states of mind of the elect in Heaven (and also Earth, as we shall see) that are the goods for which God reprobates, I suggest.

Leaving the 'main aim in creation and redemption' element of Piper's view aside, one can suggest a minimal Calvinist theodicy based on the remainder of Piper's suggestion (and on Calvin's remarks and also the *Westminster Confession's*) that is simply this: the good for the sake of which God reprobates is a richer sense of God's just character on the part of the elect, made possible by the display of God's wrath upon the occupants of Hell. There we have it. Problem solved? Is that, by itself, an adequate Calvinist theodicy of Hell?

Oliver Crisp has suggested that it is not adequate, for the following reason:

God's grace and mercy [can be] shown to all human agents in their election (in Christ), and his wrath and justice [can be] shown in the death of Christ, which atones for the sin and guilt of all fallen human agents. (2003: 137)

Crisp's contention is this: if God is after a display of his wrath, then it has already been accomplished! God has demonstrated his wrath satisfactorily in the death and suffering of Jesus. If God can elect every mere human being to salvation, and still obtain a display of his wrath, then surely God would prefer that course. So Crisp suggests that we remain just as perplexed as before as to why God would make a decree of reprobation.

One can push back against this. I don't think it is true that every good available through the display of God's wrath in reprobation is also available in God's display of wrath directed toward

Jesus in the crucifixion and in his separation from God. Interestingly, Crisp himself now appears to agree. He writes in a more recent paper:

Were Christ to be the only human person upon whom divine justice was visited, as a vicarious substitute for sinners [...], this would not have the right connection to desert because Christ does not deserve to be punished – he acts vicariously (and sinlessly) on behalf of sinful human beings deserving of punishment. (2010: 22)

So one good that is lost is the good of seeing *wrath from God that has the appropriate connection with desert*, which is surely better to see than wrath which does not. Here are some other goods which would be lacking in a world in which only Christ bore the wrath attending damnation.<sup>52</sup>

*An ongoing spectacle.* Perhaps God reprobates in order to give people in Heaven an ongoing spectacle of God's retributive justice being enacted in punishment, an aspect of the divine character that they otherwise would only be able to recall if Christ's atoning sacrifice on the cross was all of God's punishment, for Christ suffers and dies but once, and then is risen forevermore. Seeing something presently is better than merely being able to remember it, even perfectly remember it, and Christ's punishment is only of finite duration. So an ongoing perception of God's activity in wrath is a good that would be lost to the elect in Heaven.

*A better understanding of what justice demands for different sins.* If Christ is the only object of God's wrath, then all the types of sin in the world, by virtue of their imputation to Christ, are made uniform with respect to punishment. This means that in Christ's punishment we don't perceive the punishment that *different* sins deserve: we don't see the punishment appropriate to sins of greed, pride, lust, etc., we only see what happened to Christ. This limitation hampers an understanding of the justness of God, because we are hindered from seeing what justice demands in particular cases. This good looks like it would provide God with a motivation to reprobate several people with different characters all given over to different vices.

*A greater perception of the majesty of God.* Jonathan Edwards says that in the punishment of sinners God 'vindicates and honours [his majesty], and makes it appear, as it is indeed, infinite, by

---

52 I draw heavily on Jonathan Edwards in what follows. He has probably done more than anyone to provide a Calvinist theodicy of Hell: in at least four places in Edwards' *corpus* do we find substantial material which we can construe as offering reasons for God's reprobative decree. I should explain how I cite him. All of the references I give in this chapter are to the 1834 two-volume edition of *The Works of Jonathan Edwards* edited by Hickman. The first is a sermon entitled 'The Eternity of Hell Torments' or EHT (1974b [1834]: 83–89). The second is another sermon, 'The Wicked Useful in Their Destruction Only' or WUD (1974b [1834]: 125–129). The third is yet another sermon titled, 'The End of the Wicked Contemplated by the Righteous' or EWC (1974b [1834]: 207–212). The fourth is ch. III from his *Remarks on Important Theological Controversies* entitled 'Concerning the Divine Decrees in General, and Election in Particular' or CDD (1974b [1834]: 525–543). This volume is available online at: <http://www.ccel.org/ccel/edwards/works2.html>

showing that it is infinitely dreadful to contemn or offend it.’ (EHT: 87). Edwards appeals to the status principle here: the gravity of an offence should be proportional to the importance of the being against whom the offence is committed. In the eternal destruction of the wicked we discover how grave a thing it is to offend God, and by implication the majesty of his being. But if Christ alone was the object of God’s wrath then, since Christ’s punishment was only temporary and to our appearance finite, the elect wouldn’t have as full a perception of the magnitude of God’s majesty.

*Gratitude through appreciation of the nature of the alternative.* Edwards also writes of the elect in Heaven, ‘When they shall see how dreadful the anger of God is, it will make them the more prize his love. They will rejoice the more, that they are not the objects of God’s anger, but of his favour...’ (WUD: 127). This is true because of the following principle: ‘A sense of the opposite misery, in all cases, greatly increases the relish of any joy or pleasure.’ (EHT: 87). He expounds on the principle and its implications more fully here: ‘There would be no manifestation of God’s grace or true goodness, if there was no sin to be pardoned, no misery to be saved from. How much happiness soever he bestowed, his goodness would not be so much prized and admired, and the sense of it not so great, as we have elsewhere shown. We little consider how much the sense of good is heightened by the sense of evil, both moral and natural.’ (CDD: 528). So, were Christ’s misery all the elect saw, they would not have as great a realisation of the fate they had been saved from, and therefore all the joys attendant on being in Heaven would be less appreciated.

## **7.4 Calvinist Theodicy Improved**

Yet all that these various goods show is that God is motivated to create a modest number of people for Hell (a few hundred perhaps), but presumably there will be many more people in Hell than that. Traditional Christian belief has it that a sizeable percentage, if not the majority, of humanity is going to Hell. So we still face a problem: what could motivate God to reprobate such a great number? The following goods achieve this.

*Gratitude through appreciation of the likelihood of the alternative.* We have seen above that by God’s displaying eternal punishment the elect would become more grateful of their place in Heaven. We also saw that this idea doesn’t provide reason for the reprobation of a great number. But consider this quotation from Edwards: ‘When [the elect] see others, who were of the same nature, and born under the same circumstances, and plunged in such misery, and they so distinguished, O it will make them sensible how happy they are.’ (EHT: 87). The thought is this: I was just like so-and-so, yet I am exalted and they are debased, and the fact that they were just like

me makes me happier than I would otherwise be at my exaltation. But why is this? One answer concerns likelihood: it is because the closer I were to them in nature and circumstances, the greater the likelihood that I would end up like them. So when I discover that my fate has been radically different and better than theirs, my joy over my fate acquires greater intensity.

This is because your gratitude should be proportional to, in addition to the good you are the recipient of, the closeness of the possible worlds in which you fail to have it. The idea is that the closer such worlds are (in other words, the more appropriate it becomes to say ‘I might not have got it’), the greater your gratitude. Take the 2011 Tōhoku earthquake and resulting tsunami. I am (or should be) grateful that I did not die in it. But I won’t be as grateful as a Japanese person who narrowly escaped the onrushing water, in part because worlds in which he dies because of the tsunami are much closer to the actual world than worlds in which that happens to me.

So, if plucked from a sea of unbelievers, you would therefore have much more cause to be grateful. Now we see that God has reason to make it the case that the damned numerically far outstrip the elect, for if there were many people who were just like the elect were but who didn’t have faith and were damned, then this would increase the likelihood of the elect being damned considered relative to various facts, such as their being human beings, or their being born on Earth, or in New York, and so on. The more reprobated earthly companions the elect receive, the more appropriate or ‘truer’ it will be for them to say, ‘I could have been damned,’ and their gratitude at being in Heaven will increase—in proportion to both the number of these companions and the similarity of situation of these companions to themselves.

*Gratitude through appreciation of the frequency of the alternative.* This good is closely related to the preceding, but I believe it is distinct. Consider the following scenario: you attend a great banquet to which you received an invitation. The wine flows and the heart is made glad. Now suppose you discover that there are a great many people outside, all clamouring for entry, but who can’t enter because they have not been invited. Your happiness at being invited is likely to increase, and this reaction is surely appropriate. The rarer a desirable commodity, the higher it is valued. Moreover, your gratitude will increase in proportion to the number of people who can’t get in—the greater the number, the greater your gratitude.

One way of understanding this reasoning is according to the counterfactual likelihood interpretation we dealt with above: you are happier because you realise the possible-world odds were against you getting an invitation. But we don’t have to read it that way. We can interpret it as concerning only the actual world, and not what happens in others. You are happy simply because of the *actual rarity* of your position.

I am uncertain how significant the difference is between this and the previous item, but it is nevertheless clear that between them they provide the sort of motivation the Calvinist seeks. By reprobating a great number to Hell, the elect in Heaven are permitted a great gratitude not otherwise available to them: a gratitude at being part of the few that are saved.

*A greater appreciation of one's dependence upon God.* Edwards suggests that 'The misery of the damned will give them a greater sense of the distinguishing grace and love of God to them, that he should from all eternity set his love on them, and make so great a difference between them and others who are of the same species, and have deserved no worse of God than they.' (WUD: 127) The idea here, I take it, is this: when the elect observe the suffering of the damned, they will note that it is only by virtue of the grace of God that they differ from them. 'Every time they look upon the damned, it will excite in them a lively and admiring sense of the grace of God, in making them so to differ.' (EHT: 87). We have already noted the way reprobating would increase the gratitude of the elect, but here we are interested in something different. The elect will surely be drawn to contemplation of what it was that secured their salvation—what was it that gave them a fate separate from the reprobate? The Calvinist explanation lies in the decrees we discussed at the beginning. It was by God's sovereign decision that some were elected and others reprobated. Made aware that the most important aspect of their life lies in the hands of God, they are made aware of his sovereignty over them and their great dependence on him. The reprobation of many therefore draws justified attention to an important fact: your great dependence upon God's gracious decision.

*A greater appreciation of God's prerogative in salvation.* I derive this motive from the following passage from John Gill's *The Cause of God and Truth*, part 3, section 2:

[I]f God had decreed to save all men, and had prepared saving grace for all men, here would indeed have been a display of the glory of his grace and mercy; but where would have been the declaration of his wrath and justice? Especially, the glory of God's sovereignty more appears by these distinct decrees, than if no such distinction had been made; for hence it is evident, that he will have mercy on whom he will have mercy, and whom he will he hardeneth. (1855: 162)

The suggestion expressed in the first sentence has been discussed already; it is the idea in the second sentence that I wish to discuss here. Arminians and universalists are sometimes happy to grant that God is under no obligation to save anyone and that this is something important to realise (though they typically quickly add that God's good nature or character is such that it precludes him from letting any perish if he can at all help it). But if it is important to know that an agent has a right to *X*, then one good way of gaining a better understanding of this fact is to see the agent acquiring

or performing *X*. Suppose the government had a right to confiscate your property at a moment's notice and could exercise this right at a whim. This right of theirs would not be readily appreciated if they never exercised it either with regard to you or anyone else. It would be there 'on paper' but would little intrude upon anyone's experience, contemplation or decision-making. So now we have one way of reading Gill's suggestion: the 'glory of God's sovereignty' refers to his privilege to bestow and refrain from bestowing salvation as he wishes, without obligation. If God has such a privilege, such a fact is surely worth appreciating, and it is by these 'distinct decrees' of election and reprobation that a greater appreciation becomes available. This supports the reprobation of many because if God's privilege to bestow salvation were exercised too frequently then his privilege to refrain from bestowing it would be proportionately unappreciated.

*A greater appreciation of God's hatred of sin.* It is surely good for us to understand that God, because of his holiness, hates sin and wickedness. Moreover, it is surely good for us to understand that this hatred is *inexhaustible*. How could we get a better appreciation of this fact? One way is through the reprobation of many. For suppose a human judge is required to lash fifty men who had each committed a crime. He lashes, let us say, twenty of them, but after that can lash no more; not because he is physically exhausted—we can suppose that he has a very strong arm—but because his anger at their wickedness is exhausted. Absent such anger, he feels only pity. But he is wrong to give into this emotion and he should continue: the remaining thirty are no less guilty. But if God reprobates many, then it will be available for all to see for all eternity that God has a hatred of sin which is not exhausted, despite both the terrible number of the reprobate and the interminable nature of their torment. Whenever the elect look upon Hell they will be reminded that God's hatred of all that is wicked and vile remains implacable. Again our appreciation will be proportional to the number of those damned: we more clearly see that there is no worry of God's hatred petering out when we see it directed continually at a greater number.

*A greater appreciation of God's power.* A great display of God's power is also available if God reprobates many and their eternal destruction is viewed. ('The sight of the wonderful power [...] of God, manifested in the eternal punishment of ungodly men.' EHT: 87) If we suppose that an emperor only ever befriended his enemies and never waged war with them, that he wined and dined them, entertained them but never threatened them, showered them with gifts but never demanded tribute, then it is not unreasonable for the suspicion to form that the emperor is weak: he lacks the power to destroy his enemies, and that is why he always befriends them. Reprobating many gives a grand testament to God's power to destroy his enemies and frustrate their plans. As before, the more damned there are being destroyed, the greater God's power is perceived to be through their destruction.

*A justification for pragmatic preoccupation with salvation.* If Christianity is true, then the way we relate to God is the most important aspect of the human life, and acquiring salvation the most important objective. It is better if important things occupy your thoughts than unimportant things. But suppose you think you are very likely to secure the important things. In that case your attention will naturally move to less important things. But all things being equal, it is better for your mind to be occupied with the more important things. So one way in which God could secure that your mind is occupied by the more important things is by discouraging ease with regard to the important things. But suppose that all or most people are saved; this encourages ease with regard to one's standing before God insofar as it permits the following reasoning: I need not seek so hard for my own salvation as I otherwise would because I know that the odds of my being saved are in my favour. So when a man comes to learn that most are reprobated he acquires a justification for preoccupying himself with making his 'calling and election sure' that he would otherwise lack. (Thus, we see that the reprobation of many also has benefits for us in this life.)

*A justification for historical preoccupation with salvation.* The good immediately above is entirely pragmatic or this-worldly. But something similar holds from the historical perspective in the afterlife. Suppose that someone is writing the history of a wood. But the thing is everyone who entered into the wood took the westerly path at the first fork in the road; no-one took an easterly path, and consequently that entire region is unexplored. Suppose if they had taken the easterly path their journey would have been dangerous and possibly fatal. Since the path to the west offers pleasant and uneventful passage, their decision to take the westerly road would have been their most important decision. But the history of the wood would not reflect this—everyone took that path and historical attention would not be drawn to that decision as much as it would had some taken the other path. The focus of the history would instead be given to the different things that happen on the western path—things of comparatively less importance.

Now consider elect men and angels surveying the history of mankind. They know that salvation is the hinge on which a man's destiny swings, opening to him either an eternity of anguish and terror or one of everlasting joy and fellowship with God. But if all or the vast majority of these destinies swung one way, toward fellowship with God, then, because of this, salvation would become an item less worthy of historical interest. But nothing is more important in the history of men than their salvation. The reprobation of a great number makes it more appropriate for people's reflections on history to be preoccupied with the important matter of rightly relating to God than would be the case if few were reprobated.

So I have specified eight goods for which the reprobation of many is necessary: gratitude from the likelihood and frequency of damnation, justification both pragmatic and historical for

preoccupation with salvation, a greater sense of dependence upon God, of his prerogative in salvation, of his hatred of sin, and of his power. Now I discuss the question of their worthwhileness.

## 7.5 Worthwhileness and Partiality

An opponent may complain: ‘The sort of goods you suggest and the level of them just aren’t worth reprobating for. Hell is a great terror, and to reprobate so many people to such a place for the sake of such comparatively small goods as a greater appreciation of God’s attributes and so forth is not morally acceptable: the goods fail to justify because they are outweighed by the magnitude of the evil.’

Well, the theodicy I am giving here might not look good on a utilitarian analysis, but that is of little consequence: I take it few Christian philosophers are utilitarians. It isn’t clear that it fails non-consequentialist criteria for justifying goods. In this vein I shall to appeal to the view known as familial partialism, a species of what John Cottingham (1986) calls ‘philophilic partialism’. Philophilic partialism is the view that it is morally correct to favour ‘not just one’s friends, but one’s children, siblings, spouse – all who are beloved or “dear” to the agent.’ (1986: 368). Familial partialism is the view that it is morally correct to favour members of one’s own family.

What might this mean in practice? Consider the following thought experiment. A genie shows you two cages hovering above a lake of lava; one contains a family member and the other one five strangers. He tells you that you can pick one of the two cages. He will save whomever is in that cage from plunging into the fiery lake below. If you pick neither, both will fall. Are you permitted to choose the cage with the family member? A utilitarian may balk at the suggestion, but if a partiality to family members is morally expected, then it is either permissible to choose the cage with the family member, or at worst no longer clear. In more mundane cases privileging one’s family can be seen in the way you would, for example, give your child spending money rather than distributing it equally across all the children on the street, or pay special attention to your child’s grievances, an attention that you wouldn’t give to your neighbour’s child, and so on. At any rate, given that Cottingham describes philophilic partialism as ‘A pillar of all, or certainly most, viable ethical systems’ (1986: 368), it can’t be said that I am appealing to anything controversial here.

How is familial partialism relevant to the task of this chapter? I believe it is by appealing to familial partialism that the Calvinist can give a model to show how the goods given above can justify God in the reprobation of many. To explain, we can understand Christendom as containing two views on the nature of the fatherhood of God. One view, and perhaps the more common one, is

that of Universal Divine Paternity (UDP): the claim that all of humanity are God's children. But the Calvinist will probably deny UDP and hold instead to Particular Divine Paternity (PDP): the claim that it is only a significantly proper subset of humanity that are God's children, namely, the elect—those that are (or will be) Christians. God is not the Father of the persistent unbeliever, but only those he has elected to partake in the salvation of Jesus Christ. Calvin appears to hold this view:

In calling God our Father, we certainly plead the name of Christ. For with what confidence could any man call God his Father? Who would have the presumption to arrogate to himself the honour of a son of God were we not gratuitously adopted as his sons in Christ? (*Institutes*, 3:20:36)

For Calvin and those who follow him the honour of being a son of God (part of God's family) is acquired through adoption, not through creation—only believers can truthfully call God 'Father'. And of course, since it is only the non-elect that go to Hell, God is therefore not guilty of sending any of his children to Hell, and so any argument to the effect that God has violated his paternal obligation to the reprobate in reprobating him will beg the question against the Calvinist: God has no paternal obligation to the reprobate on the Calvinist view.<sup>53</sup>

So, here is how the Calvinist should explain why the comparatively lesser goods the elect receive are worth the comparatively greater evils involved in the damnation of a great number of reprobates: the elect are God's children, his family, while the reprobated are not and never were. And if we agree that familial relations are an appropriate source of privileging and partiality, then the Calvinist can say that from God's perspective, God's *first-person valuation*, the goods the elect receive, on account of his paternal relation to them, are esteemed greater, or at least equal to, the evils which befall the reprobated, despite the *intrinsic valuation* going the other way. And this is no different in principle from the way familial partialism expects you to value your own children's joy more than that of other people's.

This is why proposition (5) above is false. It was the claim that it is only *greater* goods that permit God to allow evil, but we should not be interested in whether the suggested justifying goods are *intrinsically* greater, but whether God, given the value of these things *for him*, is justified in

---

53 See also Kelly James Clark's 1995, which accepts that God is the father of all, but contends that even in that case the great differences between God the Father and earthly fathers means that God does not enter into the network of human obligations in any simple way, such that there is no easy inference from the permissions and obligations of earthly fathers to the permissions and obligations of God the Father.

permitting the named evils. The value of something for an agent can be much greater than its impersonal value.<sup>54</sup>

Can my opponent persist and claim that, even granted what I have said above, the goods still do not justify and this should be evident? I concede that I cannot *prove* that the goods the elect receive outweigh the evils to the reprobate from God's position, but I also fail to see how my opponent can prove that they do not outweigh them. I can see how he can see their comparative intrinsic value, but not how he can see that the disparity of intrinsic value is more than the paternal tie can justify. I'm afraid it isn't clear to me how much the paternal bond is capable of justifying. And I am content to leave things there. As I stated at the beginning of the chapter, I am only concerned to give a theodicy against which no *clearly probative* objection could be made, and this objection does not meet that requirement.

## 7.6 Two Objections

I deal now with two remaining objections.

*Objection from awareness by other means:* All of the goods I propose involve an item of appreciation possessed by the elect as the end which God seeks. It was suggested that God reprobates in order to give the elect an awareness of his hatred of sin, for instance. But why does God have to show us his wrath before we realise that his hatred of iniquity is inexhaustible? Why can't he, by exercising his omnipotent powers, place such an awareness in our minds directly? Or perhaps he could fill our minds with imagery which demonstrates these things? Why not, instead of actually creating people for Hell, doesn't he show the elect what it would be like if there were?

*Response:* Note first that some of the goods are immune from this criticism. You can't have a (non-misguided) gratitude at being one of the few elect if everyone is elected, for instance. Neither can the great amount of lost souls (truly) justify your preoccupation with salvation if there are no lost souls.

But this objection raises issues for the others. Let us work with an analogy. Suppose I know that Bruce is a very strong man. I have it on good authority that he can lift trucks with his bare hands. I have never seen this for myself, but nevertheless have a justified belief to this effect. However, one day I encounter Bruce at a show and there I do see him lift a truck with his bare hands. Insofar as

---

<sup>54</sup> Familial relations are one putative source of this sort of agent-relative value. By Nagel's reckoning there are three categories of value for an agent that appear irreducibly agent-relative: (i) autonomy, which includes one's own desires, projects, commitments, and personal ties; (ii) deontology, which includes respecting rights—in general the idea that one must not be a doer of certain things; (iii) obligation, which includes special obligation one owes to spouses, children and other family members (1986: 165).

my appreciation of Bruce's strength goes, I judge it much enhanced by witnessing this exhibition. What I gained there was not knowledge—I already had a true, justified belief that Bruce was a very strong man who could lift trucks—what I acquired was something like a feeling of awe at Bruce's strength.

Could God provide this without Bruce lifting anything? I agree that God could do this, but then the awe would be irrational—it would lack justification (or whatever the analogous relation is that governs emotions). To return to the example of the Japanese man who narrowly escapes the 2011 tsunami, God could make me feel, as I sit at my desk, as relieved and euphoric as he did, but it would be irrational because there is nothing in my environment to prompt it.<sup>55</sup>

But perhaps God could provide a mental cinematic showcasing Bruce's strength. Yet one feels this is inadequate for the sort of reason Nozick gave with his 'experience machine' thought experiment (1974: 42–45). Climbing a mountain in real life is a better thing to desire than climbing it in an experience machine, even if the experience is identical. Perceptive experience plus reality is greater than merely the experience, and therefore it is better for the elect that they receive the former. Also, once you realise that the display a correct representation of reality, a great deal of the sting is appropriately lost: a man who reacts emotionally to everything he sees in a film as if it were real is dysfunctional. I conclude that this irreducible advantage of displays that truly relate reality to us provides God with significant motivation to give real demonstrations of his attributes, as opposed to mere cinema.<sup>56</sup>

*Objection from God's love:* It might be thought that the theodicy I give here fails to do justice in some way to God's love. Recall Baggett and Walls's complaint from the start of the chapter that reprobation was a 'paradigmatic example of hateful behaviour, not loving behaviour.' (2011: 74).

*Response:* Hateful behaviour is behaviour that is motivated by hate. So if we can consistently suppose that God in reprobating is doing so from a loving motive, then this objection will lose its force. But I think it is clear that on the theodicy I have sketched here God's motive is a loving motive, for he reprobates out of a desire for his creatures' good, albeit the good of the elect. Here is how I see the explanatory order of God's decisions on the account I have provided: God first sets his love on a particular set of possible creatures; these are his elect; he loves them as a father, and desires what is good for them. To bring this love to consummation he decides to create them. But what is good, indeed best, for the elect is that they love God and understand his nature and relation

---

55 I think this is a neglected evil in the use of drugs. They make you happy, but they often make you happy without justification.

56 Kyle Scott drew my attention to the helpful case of couples in love: they, despite knowing that they love each other, often buy their partner gifts as a means of demonstrating that love. This helps illustrate the general principle that it is better to see that *p* rather than merely to know that *p*. Suppose a suitor, instead of buying a ring, bought a picture of him doing so! I do not think that would be well-received.

to them. Now God can partially achieve these objectives by displaying his power, his hatred of sin, his sovereignty in salvation, and all the other items I listed previously. But God cannot use the elect for these displays because that would be inconsistent with his paternal love for them. So God also forms the intention to create other human beings (the reprobates) for whom he will not have a paternal love in order to manifest these displays. But the important point here is that it is plain to see how God's love for certain of his creatures is the overarching motive here, illustrated by the fact that the decree of reprobation is explanatorily posterior to the decree of election—the reprobate is useful instrumentally, for the goods which God, from a loving motive, desires to bestow on the elect.<sup>57</sup>

## 7.7 Conclusion

That concludes my Calvinistic theodicy of Hell. Between them, this chapter and the preceding one provide a comprehensive theodicy of Hell: they give an account of the justice of Hell, and also explain why it is that God would be motivated to arrange matters so some people would end up there. And all of it consistent with Calvinist assumptions. Since it is worries about the goodness of God in relation to Hell and reprobation that appear to be the chief concern amongst Calvinism's critics, I take this material to be a productive step forward in the debate.

One might worry that the reliance of the theodicy on the denial of universal divine paternity is a weak point of the theory. For that reason the following chapter will be devoted to the issue of the fatherhood of God—its object being to put particular divine paternity on a secure philosophical footing.

---

<sup>57</sup> This also makes it clear that the theodicy I am offering is supralapsarian.

## Chapter 8—Divine Love and Paternity

### 8.1 Introduction

Many Christians may have a knee-jerk reaction against the suggestion that God is the Father only of the elect, or only of Christian believers. And there are some challenges to the idea, both Scriptural and philosophical. I intend to deal with those challenges here, but first I will make a *positive* case for Particular Divine Paternity (PDP), because that will set the stage for a proper assessment of the objections to it. Thomas Talbott is a leading defender of the idea that God is the father of all, Universal Divine Paternity (UDP), and I will discuss his material below. His exchange with Jeff Jordan on closely related matters forms an excellent entry point to the debate.

### 8.2 The Talbott-Jordan Exchange

There was a recent, and fascinating, exchange between Thomas Talbott and Jeff Jordan in the past few issues of *Faith and Philosophy*. The debate centred around the following proposition:

(L) If God exists and is perfect, then God's love must be maximally extended and equally intense. (2012: 53)

Jeff Jordan argued that (L) is false. His argument, in essence, was this: there are no defects in God; a life without deep attachment is deficient; ergo, God has deep attachment. But deep attachment implies an inequality of love; therefore, God's love for all is not equally intense. He draws one significant implication (of obvious relevance to our purposes) from this toward the end of his paper—he argues that the falsity of (P) follows from the falsity of (L):

(P) If God exists, then the relation between creator and human is that of loving parent and child.

In short, he argues that God's deep attachments imply that he is not the Father of all. Why does this follow? Because 'Implicit in (P) is the [equal-intensity] requirement, as a good and loving parent

loves his children equally. But with strong reason to deny (L), we also have strong reason to deny (P).’ (2012: 68). That is the connection of chief interest here.

Talbott responded to Jordan in his ‘The Topography of Divine Love: A Response to Jeff Jordan’ (2013), making a number of objections. Jordan offered a brief rejoinder to Talbott in his ‘The Topography of Divine Love: A Reply to Thomas Talbott’ (2015).

What I shall initially do in this chapter is provide commentary on and criticism of the Jordan-Talbott debate. Although I offer criticisms of both parties, I ultimately endorse Jordan’s overall position. I think that Jordan’s argument is wanting at different points, and I offer two different reconstructions of Jordan’s argument that avoid the problems I believe are present. Furthermore, I take Jordan’s argument one step further, and to argue that the falsity of (L) not merely shows that God is not the father of all, but also that soteriological universalism is false, and therefore that not all will be saved.

### 8.3 Flatness in Love

Let us begin with (L) itself. Here it is as it is in Jordan’s initial article:

(L) If God exists and is perfect, then God’s love must be maximally extended and equally intense. (2012: 53)

The idea here is that the topography of God’s love is *flat*. It is not higher for some individuals than it is for others—it extends to everyone and it is the same for everyone.

Quite aside from any issues involving deep attachment, we should take issue with (L). (L) doesn’t specify those objects to which God is directing his love. Suppose it includes absolutely everything. But it would be a mistake, I take it, to say that God loves rocks and pebbles with the same intensity with which he loves human beings. The same point can be made if we restrict the love in question to living things: surely, because we are more important than sparrows (Matt. 10:31), it is fair to say that God has a greater love for us than for sparrows. Even if we restrict the love in question to persons, counterexamples continue to arise. Consider angels. They are persons. It is plausible to suppose they are greater beings than we. If that supposition is correct, then we would expect God to love angels more than us. To avoid these complications, it might be thought best to restrict the principle to human beings (I assume every human being is a person):

(L') God's love to human beings is maximally extended and equally intense.

Are there any obvious problems with (L')? One concern is that it appears that Jordan wished that more is to be considered as part of (L') than he actually expressed in it. He writes, 'The flatness requirement should be understood to require not just equality but also maximal intensity—every human is loved by God to the same significant degree.' (2012: 53). It appears that Jordan wishes the claim he is attacking to imply not merely that God's love to human beings is equally intense, but also maximally intense. In that case, we can capture what was intended by (L) as follows:

(L'') God's love to human beings is maximally extended and maximally and equally intense.

Is this a plausible enough principle with which to proceed? It is not. To claim that God is maximising his love for every human being is unreasonable. For one thing, we are surely to suppose that God loves himself above all his creatures because he perceives that he is the being most worthy of honour and love. This can't be so if God loves us as much as he loves himself. Or suppose there is a race of beings, about which we have no knowledge, that far excels us in natural grandeur and greatness of being. Surely, given their natural greatness of being, it would be appropriate for God to love those beings with a greater love than that with which he loves us human beings. It would be a *defect* for God to love us with the same amount of love as that which he gives to this other race of beings, for then God would not be sensitive to the difference in greatness of being between us and them. But if God loves every human being maximally, that is to say, with as much love as he can muster, then God would love us at least as much as he would any beings greater than us. But that is the wrong result. We should therefore amend the maximality requirement to include reference to what is reasonable given the nature of the being loved:

(L''') God's love to human beings is maximally extended, equally intense and as maximally intense as is reasonably possible relative to the nature of human beings.

Is this adequate? Some issues remain. Consider God incarnate in Jesus. He loved some people more than others. The gospel of John speaks of the 'disciple whom Jesus loved' (John 13:23; 19:26–27), and the plain implication here is that this disciple held a special place in Jesus's affections, that Jesus delighted in his company and prized his friendship above that of the other disciples. So, even if we grant that God the Father's love for all human beings is equally intense, nevertheless, when we add this special love Jesus had for the beloved disciple into the mix, then we will get the

conclusion that (L''') is false because the beloved disciple will be loved in this special way by Jesus in addition to whatever love comes from God in other ways.

I think that the case of the beloved disciple is indeed a *prima facie* reason to be suspicious of the equality requirement in claims like (L'''), but I will not amend (L''') to accommodate the problem, because I think the defenders of (L''') could work their way around it. This in two ways. First, they could deny that the love would be summative in the required fashion. It may be, if God's love for every human being is infinite, that the special love Jesus had for the beloved disciple is 'swallowed up' in the infinite love God has for all people in other respects, such that the love God has for the beloved disciple remains no greater than the love God has for anyone else. Second, they might point out that God the Father could always ensure that the love be 'fair' in the relevant sense by, say, reducing the love that he has for the beloved disciple to the same extent that Jesus was fond of him. In this way God could ensure that his love panned out in the required egalitarian fashion.

Ross Parker (2013) has pointed out one last problem. There is a tension between the equality and maximality requirements on intensity. (L''') as it stands requires a strict equality of intensity. In other words, God's love for all human beings is supposed to be as strong as it reasonably can be *given* the equality requirement. Such a requirement leaves open the possibility that God wants to reasonably love some human beings more than others (perhaps they are naturally more deserving), but he is stopped by the equality requirement. Another approach would be to privilege maximality. Claim that God's love for mankind must be such that, for every individual human being, God loves that human being with the greatest, the maximal, amount of love he could reasonably have for that human being. This may lead to an inequality in intensity of love, for it might be that loving every human being to the maximum would involve some people being loved more than others, for some people's maximums may be greater than other people's. Yet equality would still be preserved in a sense. God's approach would be equal: he would try to love each person as strongly as he possibly reasonably could.

Surprisingly, Ross Parker claims that in personal correspondence with Jordan (2013: 445, n. 6) Jordan claimed that it was this latter sense, which privileges maximality for each individual, that he intended. I find this surprising because not only does it render the frequent talk of equality in love and the 'flatness' terminology, which Jordan was so fond of, ancillary, but Jordan himself distinguishes between loving equally and loving fully—loving equally and loving as much as the recipient is able to take—in his article. He says that 'Perhaps it is true that God loves fully; nonetheless, there's good reason to deny that God loves equally.' (2012: 67). For that reason I will not amend (L''') to remove the equality requirement. It was precisely the equality requirement that Jordan was taking aim at—he was not concerned to deny that God loved all human beings fully—

and therefore the equality requirement remains essential to the discussion. So, it is (L''') that I will work with and take aim at in what follows.

## 8.4 Deep Attachment

So, what is Jordan's argument against (L''')? A premised argument is as follows:

- (1) If  $x$  is a deficiency in a human being, then God does not possess  $x$ .
- (2) A human being without any deep attachment is deficient.
- (3) Therefore, God has deep attachment.
- (4) Therefore, God's love for human beings is not equally intense.

And it follows from (4) that (L''') is false. What might we say about this argument? Much requires explanation. (1) is an expression of an inverted approach to perfect-being theology. Just as we are inclined to ascribe to him what makes us great, so we are inclined to deny to him what makes us bad. Such a principle is not, however, above challenge. It isn't clear to me why God's nature might not take on some things intrinsically imperfect if they enable the sum total of God's greatness to be greater than it otherwise would have been.

Jordan himself admits that a principle like (1) is hard to establish, and he gives a lengthy defense of it. I don't wish to engage in any prolonged discussion of the principle, however. I shall merely note two things here. First, that, as Parker (2013) notes, as Jordan defends his argument it turns out that it is quite possible to marshal an argument against (L''') that doesn't rely on this premise. Second, we can, if we wish, replace (1) with (1')

(1') If  $x$  is a deficiency in a human being, then we should be *ceteris paribus* inclined to believe that God does not possess  $x$ .

Such a principle would get us the weaker conclusion that we should be inclined to believe, other things being equal, that (L''') is false. This, I think, would still be enough to motivate Jordan's and my position that God is not the Father of all, and to shift the burden of proof to the opposition.

Premise (2) contends that a human life devoid of deep attachment is seriously defective. What is deep attachment? Deep attachment is the sort of thing one bears to one's friends, family and romantic partner. Jordan puts it this way:

[C]onsider that a human who loved all other humans equally and impartially would have a life significantly impoverished. Much of the richness of life flows from one's friendships and one's spouse and one's children, and within these attachments there is a love which is neither impartial nor equally shared by all other persons, as one loves her beloved more than she does others. It is not just that one manifests her love for the beloved differently from how one manifests her love for others. No, a person appropriately loves his own children more than other children. And without the inequality of love, one's life would be diminished. (2012: 60–61)

Jordan contends that an inequality of love is the necessary result of deep attachment. But because God's perfection demands deep attachment, it therefore demands inequality of love. (L''') is therefore false, because (L''') requires equality of love.

Towards the end of his paper, as I mentioned, Jordan draws the connection with divine paternity. He notes that many people hold to the following assumption:

(P) If God exists, then the relation between creator and human is that of loving parent and child.

Atheists as well as theists often assume (P).<sup>58</sup> But Jordan contends that this proposition is also false. Why so? Because it is part of the duty incumbent on a parent to love all their children equally. To show favouritism amongst one's children is an intuitively wicked thing to do.<sup>59</sup> We can therefore extend the argument given in (1)–(4) as follows:

(1) If  $x$  is a deficiency in a human being, then God does not possess  $x$ .

(2) A human being without any deep attachment is deficient.

(3) Therefore, God has deep attachment.

(4) Therefore, God's love for human beings is not equally intense.

(5) If God were the father of all human beings, then his love for all human beings would be equally intense.

(6) Therefore, God is not the father of all human beings.

Some might object to (5) on the ground that love is a feeling and that it would be absurd to suppose it is part of the duty of a father to feel equally pleased or delighted in the company of each of his

---

58 Jordan points us to Rowe's famous 1979 paper as an instance of atheist adherence to (P).

59 It is also tacitly condemned by Scripture. The favouritism that Rachel shows to Jacob, and Jacob to Joseph, and that Elkanah shows to Hannah leads in each case to familial strife.

children. Some children might have personalities or traits of character that particularly enamour their father, and that elevate them in his felt affections. Surely it is not wicked for the father to acknowledge this and do nothing to mitigate the inequality in feeling (provided, perhaps, that the inequality is not too extreme). But if that point is conceded, then why should equality in love be expected of fathers, and of God in his role as father?

I agree that such inequality of feeling (provided it is not terribly great) doesn't imply a failure in paternity, but I am not persuaded that such feelings are part of love, or at least the sort of love that we are considering here in relation to fatherhood. The account of love that Jordan relies on is drawn from Frankfurt, and it is an account I think is broadly correct. Here is Frankfurt on love:

[A] lover identifies himself with what he loves. In virtue of this identification, protecting the interests of his beloved is necessarily among the lover's own interests. The interests of his beloved are not actually other than his at all. They are his interests too. (2004: 61)

Later on he notes that love

consists most basically in a disinterested concern for the well-being or flourishing of the person who is loved. (2004: 79)

We can combine these two elements, as Jordan does, and say that to love A is to identify with A's interests, where an interest is either (i) something that is good for A, or (ii) something A desires that is not bad for A. But such an account, if true, makes it clear that a father's delight in one of his children's antics over another one of his children's is no inequality in love, because it is compossible with his identifying with both of their interests. He might delight in their behaviour and personality to different degrees, but just so long as he desires to the same extent that their interests are satisfied, then there is no inequality in love.<sup>60</sup>

It might be further pressed that it is very hard, in ordinary circumstances, for a father even to equally desire the satisfaction of his children's interests. I agree with this, and to the extent that it is hard, this diminishes the father's guilt for his failure to love equally. But this casts no doubt on (5), for God will satisfy the paternal ideal perfectly. Although a man might not love his children equally on account of the incapacity of the flesh and unstable human affection, and be blameless for the inequality in love for that reason, it is nevertheless obvious that in so doing he would have departed from the *ideal* of fatherhood, and that it would be better for him to love equally. But God is surely

---

<sup>60</sup> In a similar vein, Brian Leftow, in an unpublished lecture entitled 'Election', makes the suggestion that it is part of the duty of the father to *love* his children equally, but not to *like* them equally.

the perfect and ideal father, the father above all fathers, and will therefore love all his children equally, if that is what fatherhood requires.

## 8.5 Does Deep Attachment Require Inequality in Love?

Having laid out the case and warded off *prima facie* objections, I turn to more substantial objections. The big complaint with the argument as it stands is this: why should we accept the move from (3) to (4)? Why should it follow from the fact that God has deep attachment that God has any inequality of love?

Jordan needs more than merely the claim that

(A) God has deep attachment.

He also needs the claim that

(B) God does not have deep attachment to all human beings.

For while it is plain that deep attachment does demand that the deeply attached are loved more than the non-deeply attached, it is not plain that God must, if he is deeply attached, make sure he is not deeply attached to all. But it is only this latter claim, (B), that would guarantee the desired conclusion that God does not love all equally.

Talbott and Parker both make this same complaint. In our experience, having a deep attachment to someone does involve an increase in intensity of love for that person, and not to everyone else, but it seems plausible that the good of such relationships doesn't depend on the increase in affection being an increase *relative to everyone else*, merely upon its being an increase. So, why can't God have this increased intensity of love towards everyone?

Of course, in our case, for us to befriend the whole world would be emotionally impossible. Our time, resources, and emotional capacity are finite, and for us to try to have a deep attachment to all human beings would inevitably result in a cheapening or watering down of our attachments. But no such problems would beset God. His resources, emotional and otherwise, are without limit, and there is no risk of his being emotionally overtaxed in loving all.

So, why should we believe that God's having a deep attachment to some precludes his having a deep attachment to all?

## 8.6 Romance and the Church

Jordan gives his own argument, and I will discuss that presently. But first I would like to offer my own suggestion. Jordan named three types of deep attachment: friendship, family, and romance. When it comes to friendship and family, I am inclined to agree that there is no obstacle, from the relations themselves, to prevent those relations from being extended indefinitely. It would not be inconsistent with, nor would it undermine, my friendship with X should I take up a friendship with Y. It doesn't appear to be in the nature of friendship that it prevents the befriender from extending it to more and more people. The same holds for family; there doesn't appear to be any problem with a couple having one more child and thus extending their family even further. That wouldn't, by itself, undermine the already existing familial relations.

But with romance things appear different. There we do seem to encounter a pressure from the nature of romantic love that demands a certain exclusivity—the insistence that the relation should not be extended indefinitely, but remain restricted to lover and beloved. I think we can therefore ask with some profit whether or not God is to be considered as involved in romance. For suppose God is. In that case, the object of God's romantic affection would be the receiver of a special and intense love, and that love would, by its nature, prohibit God from spreading it to all.

One immediate riposte to this suggestion might be to appeal to intra-Trinitarian relations. Yes, God has exclusive, romantic affections, but they obtain only between the Father, Son, and Holy Spirit, and are not borne to any of humankind.

I think the best response to this riposte is Scriptural. The restriction of divine romance solely to the members of the Trinity is at odds with the biblical portrait. Consider the following verses:

“Therefore a man shall leave his father and mother and hold fast to his wife, and the two shall become one flesh.” This mystery is profound, and I am saying that it refers to Christ and the church. (Eph. 5:31-32)

And I saw the holy city, new Jerusalem, coming down out of heaven from God, prepared as a bride adorned for her husband. (Rev. 21:2)

Both Paul in *Ephesians* and John in *Revelation* are happy to describe the relation that Christ bears to the Church as being that of bridegroom to bride. This is enough to make plausible the claim that

(C) God has a romantic attachment to his church.

Is (C) enough to get us to (4), to the claim that God doesn't love all equally? That depends on whether one is a *soteriological* particularist or a universalist. Are only some human beings saved, or are all saved? If the former, then, yes, I think it does. Because one can supply the particularist premise and contend that

(D) The church, ultimately speaking, doesn't include all human beings.

And (C) and (D) together imply that God loves those human beings who are part of his church with a greater love than those outside it.

What if one is a universalist? In such a case, (D) would be denied: all people are ultimately included in the church, because all are ultimately saved. This would block the argument because even if it is granted that God has a deep romance-like attachment to the church, the universalist insistence that the church extends to all blocks any move one might wish to make to the effect that God does not bear that special love to all.

Because I intend to give an argument for particular divine paternity that should be persuasive to Christians of every stripe, the universalist as well as the particularist, I shall move on. Although, given that most Christians are soteriological particularists, the realisation that there is a plausible argument from that premise to the denial of (L'') should not to be ignored. If my joining the church, Christ's bride, doesn't increase God's love for me relative to those outside, then what will?

## 8.7 Argument from Incompatible Interests

Here is the argument that Jordan gives to bridge the gap between God's having deep attachments and God's having an unequal love for all—to move from (3) to (4). Jordan, as we noted, claims (plausibly) that to love someone is to identify with their interests. But for God to love all equally, says Jordan, would be for him to identify with incompatible interests, something a perfectly rational being, such as God, could not do. Here is one example of incompatible interests that Jordan gives:

Suppose you have an ample supply of tickets to an event, which both Smith and Jones greatly desire to attend. But Smith will attend only if Jones does not. Although you prefer going with both, you

decide to attend with Jones even though you know this means Smith will not attend. To secure the interests of one may entail thwarting those of another. (2012: 62)

Here is another:

Consider Jones, a participant in a scholarship pageant in which only one contestant will receive a full scholarship to a university—a scholarship vital to Jones’s future. ... It is obvious that among the best interests any of us have are some compossible with those had by all others. But it is also clear that some are not—Jones’s best interest in winning the pageant is not compatible with the best interests of the other contestants as it is in each of their best interests to win the pageant also. If there are zero-sum situations of any sort the winning of which is among the best interests of more than one person, then conflict among the best interests of persons is not just possible, but unavoidable. (2015: 185)

It therefore seems plausible, especially in the light of the latter example (in the former case, one is suspicious that Smith’s desire is wicked and should not be encouraged), that there are indeed incompatible best interests of persons. What follows from that? At this point Jordan offers two ways of pushing his argument forward. Unfortunately, I don’t think either way is successful. The first way is based on rational norms, and he puts it as follows:

Since persons have incompatible interests, it follows that one cannot befriend all in the deepest sense, as it is not possible to identify with the interests of all, when those interests are incompatible. [...] In other words, even God cannot love or befriend every human in the deepest way. With this result, it is clear that the obstacle against universal friendship of a deep sort is not just a practical matter, but an in-principle matter. (2012: 62–63)

But this argument is in error. We noted before that Jordan’s account of S’s interest was either something good for S or something S desired that was not bad. To identify with S’s interest in acquiring A is surely nothing more than *desiring* that S acquires A. But it is perfectly possible to have incompatible desires. A maiden may be courted by two paramours, and find her heart truly drawn to both of them. But she cannot marry both; ergo, she must choose one. But it doesn’t follow from her choosing to marry one of her suitors that she had no desire to marry the other. She desired to marry each of them. For the same reason, it doesn’t follow from the fact that God is forced, by acting on one person’s interest, to frustrate another person’s interest, that God did not desire to satisfy the latter person’s interest.

But Jordan is alert to a criticism of this nature. He writes, ‘One might, however, object that the connection between the degree of love and identifying with the interests of the beloved has been drawn too tight. ... [F]ailing to satisfy an interest, one might object, does not imply that one does not fully identify with it. So, the account of love on which the argument rests is faulty, as one can identify with another’s interests without always satisfying that interest. (2012: 64–65).

What is his response? He gives two responses, but both are unsuccessful. The first is this:

It is certainly true that one can fail to satisfy an interest which one in fact identifies with, but it is also true that it would be tendentious to claim that one could fail to satisfy an interest when one could and yet one still identifies with that interest, since one will seek to promote those interests one identifies with. (2012: 65)

But it is false to claim that satisfying an interest when one can is necessary for identifying with that interest. Take a Lifeguard Situation. The lifeguard sees two children drowning, but the distance between them is great. He knows he can save one but not both and, accordingly, saves one of them. It seems false to say that he didn’t identify with the interests of the child he left to drown because he could have saved them but did not. He did identify, but was caught between a rock and a hard place. I think that what does follow from one’s desiring  $x$  is not that one will do or acquire  $x$  when one can, but that one will do or acquire  $x$  when one can *other things being equal*. In other words, that one is disposed to acquire or do  $x$ . But, of course, a disposition can be possessed in absence of its manifestation.

The second response Jordan gives is this:

No one accepts that one can believe a known contradiction, or at least, no one should accept that a known contradiction can be believed, as one cannot believe something true which she knows is false. This is true even though all of us no doubt believe many undetected inconsistent propositions. For a similar reason, no one could identify with (take as his own) interests which are known to be incompatible. ... This reasoning applies *a fortiori* to a supremely rational being. (2012: 65)

But this reasoning is similarly weak. If there is nothing irrational about having incompatible desires (and there surely is not, given cases such as the Lifeguard Situation described above), then there is nothing irrational about identifying with incompatible interests, because coming to identify is merely coming to have certain desires. What would be irrational is *intending* to satisfy two

incompatible interests. *That* is evidently something a perfectly wise and rational being would not do.

## 8.8 Argument from Expressively Successful Deep Attachments

Jordan's overall argument therefore grinds to a halt at this juncture. We need a way forward, and I hope to provide one.

For the sake of completeness, I should mention the following suggestion. One might think that one way of fixing Jordan's immediately preceding argument would be to drop the claim that it is irrational to identify with incompatible interests, and focus instead on the weaker claim that one shouldn't identify with incompatible interests *insofar as it is up to one*. Prudence recommends such a policy. If one has incompatible interests—that is to say, incompatible desires—then this can lead to conflict and disappointment in the satisfaction of one's desires. Better to avoid that sort of thing if one can.

But this weaker claim, while probably true, is too weak to get Jordan what he wants. It might be rejoined that the goodness of God's nature demands that his love overflow in equal measure to all human beings, or that God has a duty to his creation to love all people equally. So, a bridge from it being prudential to avoid incompatible interests wouldn't imply that God does not have incompatible interests.

I think the best way to push Jordan's argument forward at this juncture is by focusing on what I will call 'expressively successful' deep attachments. We can understand better what an expressively successful deep attachment is when we contemplate the following paradigm case of an expressively unsuccessful deep attachment. Consider the story of Cosmo von Wehrstahl in George MacDonald's *Phantastes*. A strange mirror comes into Cosmo's hands through which, every night at 6 o'clock, he sees a beautiful woman enter his apartment. She lies down on his couch and goes to sleep. She is present in the reflection of Cosmo's apartment in the mirror, but not in the apartment itself. Cosmo can see her, but she cannot see Cosmo. Cosmo is quickly enamoured of her beauty, but frustrated at the distance betwixt him and her. MacDonald writes,

Meantime, how fared Cosmo? As might be expected in one of his temperament, his interest had blossomed into love, and his love ... into passion. But, alas! he loved a shadow. He could not come near her, could not speak to her, could not hear a sound from those sweet lips, to which his longing eyes would cling like bees to their honey-fountains. (2008 [1858]: 160)

There is no reason, we can suppose, to doubt the sincerity and genuineness of Cosmo's love. But the relationship he bears to the object of his love is nevertheless seriously deficient. He cannot effectively express or act on his love for her. There is no possibility of consummation. Cosmo has a deep attachment, but it is an expressively unsuccessful one. To have a deep attachment that is expressively unsuccessful is a big problem and can greatly disorder the affections (indeed, Cosmo pines away on account of his unrequited love). An expressively unsuccessful attachment is not, therefore, something that we should rush to posit of God. So, let us say that it is not only the presence of deep attachments that is necessary for a non-deficient life, but the *absence of expressively unsuccessful ones*.

But now I think we do have a good argument against the claim that God has a deep attachment to all. For suppose he does. In that case those deep attachments are expressively infecund in all cases where incompatible interests are involved. Something as simple as sport will reveal the problem. Consider two sports teams having a match. If God is deeply attached to every player on the pitch, then I don't think he can support one team over the other, where supporting involves some sort of emotional investment. Because to support one team over the other would be for him to betray the deep attachment he has to the other team. If God loves all equally and deeply, then he is prevented from having an emotional investment in one person or set of persons when their interests conflict with those of another person or set of persons. This is to place something of a stranglehold on God's emotional life. But if we suppose that God loves unequally, then there is no problem. God can support one person or set of persons over another on account of an increased love for the first person or set of persons. Just as one will justify supporting one team rather than another because one has a friend on the first team, so God's supporting one side can be explicable through his increased love for that side.

But God's supporting and investing in certain people or groups of people is probably best understood, not in relation to sport—as a spectator watching an event with an uncertain outcome—but to something akin to writing a novel. An author will pen his different characters and typically warm to the hero more so than to the villain. He will relish in the villain's eventual downfall and exult alongside the hero in his victory. But if God, as the author of the story of the universe, loves all his characters equally and deeply, then he is prevented from exulting in the victory of one when such a victory necessarily involves the defeat or frustration of the interests of another.

For an illustration of how bad it is for someone deeply attached to two different people who are pitted against one another, we might look to the Klitschko brothers. These two gentlemen dominated boxing for the past decade or more, but their mother made them promise not to fight one

another professionally—that would be a spectacle she could not bear to see—her own children fighting.<sup>61</sup> I find it very unsatisfactory to think of God as in the same predicament as the Klitschkos' mother in most or even many cases of conflicting interests. If we are all God's children, then every conflict human beings have with one another is a conflict God has to watch between his own children—it is always within the family.

If it is protested that sports are really a trifling matter and that on any important consideration or affair of great moment it will not be the case that God suffers from any expressive stifling, I think that such protests are mistaken. Consider wars. They are affairs of great seriousness, but again God is not permitted, if he is deeply attached to all, to be on one country's side rather than on the other (assuming that one country's conduct is not markedly wickeder than the other)—would a parent whose children were sadly determined to spill one another's blood be permitted to side with some? Surely not. This is a striking result given how often countries have gone to war believing they had God on their side. But if God is deeply attached to all, then it looks as if God's attitude is best understood as him washing his hands of the whole affair, supportively speaking. And if wars raise any complications on account of the great number of people involved and the variety of their situation, we might turn to duels, or any form of lethal single combat. In those cases there are only two individuals, who may both be fighting out of honourable motives, yet God, if he is father of all, cannot invest himself in the success of one of the combatants, one of his children, over the other.

## 8.9 Summarising the Case Against UDP

Let us take stock. I have offered two arguments for the claim that God does not love all equally. The first was an argument from the Scriptural presentation of God as involved in a romance with the church. If that is so, then God will surely love the members of the church more than non-members. This argument, it was conceded, requires soteriological particularism.

The second argument concerned expressively successful deep attachments. (L'') required that God loved all human beings equally and to the maximum degree rationally possible and consistent with that equality requirement. So (L''') implies that God is deeply attached to all human beings. But if God is deeply attached to all human beings, then we find the expressive successfulness of those relationships challenged. Because it is obvious, as a contingent fact, that many people have interests opposed to those of other people, this means that God is prevented from celebrating or otherwise emotionally investing himself in the success of one of the people he is deeply attached to

---

61 See: <<https://heavy.com/sports/2015/04/klitschko-brothers-net-worth-who-wins/>>, accessed 28 Sept, 2019.

over the other. Because we don't want to think of God as emotionally hamstrung or stymied in this way, we should reject the claim that God loves all equally, and therefore reject (L'''). God is not deeply attached to all, and neither does he love equally.

But either conclusion—that God is not deeply attached to all, or that God does not love all equally—appears to show that God cannot be the father of all, because fatherhood demands all children be loved equally. So, from these considerations, Particular Divine Paternity is well-motivated, and Universal Divine Paternity is not. We don't need Calvinist theology, therefore, to motivate the denial of the universal fatherhood of God—it is something every Christian should be sympathetic to.

## **8.10 The Case Against PDP: Creation and Paternity**

Now I turn the objections that can be brought against PDP, and the arguments in favour of UDP. Talbott believes he can sidestep the arguments Jordan gives, and give a direct proof that (L''') is true. He asks us to consider the following property:

Consider a property that one exemplifies only when one's love extends maximally and with equal intensity to every person that one freely chooses to bring into being (2013: 304)

He calls this property maximally extended parental love (PL). Talbott suggests that a being that failed to love equally those he freely chose to bring into being, failed to exemplify PL, would be deficient. So God loves all human beings equally, because we shouldn't attribute deficiencies to God (other things being equal). He concludes, 'a morally perfect God would, of necessity, love (or will the best for) each of those persons whom he freely chooses to create.' (2013: 304). Thus, Jordan's argument, in Talbott's eyes, can be countered with his own, so even if there is no available response to Jordan's argument, Jordan's argument for his conclusion can be matched by an equally good argument for the opposite conclusion.

So, what of Talbott's argument? Is it true that one who fails to love with equal intensity all those persons he freely chooses to bring into being is morally deficient? It looks true that one should love all one's children equally. Yet I don't think that every person you freely choose to bring into being should be considered your child, nor that you are necessarily deficient for not loving them as much as some other people you freely chose to bring into being.

The intuitions Talbott wants to appeal to are weak in cases where the offspring you create are fully grown. Consider the following thought experiment. Suppose you are introduced to futuristic six-foot capsule with a button on the side. If you press the button, a fully grown adult male will be formed in the capsule and emerge from it. You decides to press the button. The man emerges from the capsule, thanks you warmly for creating him, and declares that he would like to be on his way. You can't really stop him, so off he goes. Are you guilty for not loving this man as much as your own biological children? Intuitively not, but then there is something wrong with Talbott's principle.

But what if the principle is restricted to the creation of human infants or babies? Does Talbott's principle hold good then? Again, I don't think so. Suppose that, in the year 3000, if a couple can't have a baby naturally, then can go to their local caretaker, who can create, using advanced science, day-old babies in his laboratory. It is the caretaker's job to create babies and to give them to couples who need them. Are we forced to view the creator of these babies, the caretaker, as deficient for not loving the babies he creates in the laboratory as much as the babies he creates with his wife? I don't feel any pressure to say that. And these two thought experiments are good counterexamples to the suggestion that creation is sufficient for paternity.

I think more plausible variation on PL that is much harder to counterexample, would be this: '[the] property that one exemplifies only when one's love extends maximally and with equal intensity to every person that one freely chooses to bring into being by the *normal biological process of sexual reproduction*.' That would deal with both of the above scenarios, because neither capsule nor caretaker are creating persons in the normal biological manner. However, how are we to transfer the principle over to the divine case? Is there anything analogous to biological reproduction in the case of God? I think there is, but it won't sit well with Talbott and his ilk.

Consider the Scriptural texts that speak of salvific adoption or birthing. Here is a sample:

Truly, truly, I say to you, unless one is **born** of water and the Spirit, he cannot enter the kingdom of God. That which is **born** of the flesh is flesh, and that which is **born** of the Spirit is spirit. Do not marvel that I said to you, 'You must be **born** again.' (John 3:5–7)

Of his own will **begat** he us with the word of truth, that we should be a kind of firstfruits of his creatures. (James 1:18, AV)

And now, **little children**, abide in him; that, when he shall appear, we may have confidence, and not be ashamed before him at his coming. If ye know that he is righteous, ye know that every one that doeth righteousness is **born** of him. (1 John 2:28–29, AV)

The Holy Spirit's regenerating, saving operation in man is described in these passages and others as a birthing. Becoming a Christian is to be born again, not, like your initial birth, through the operation of an earthly father, but this time through a heavenly one: God. I'd say that if there is anything analogous to biological reproduction in God's dealings with man, the Scriptures give good indication that it is the Holy Spirit's work in regeneration that counts as God's giving birth. But that would imply that only believers are God's 'biological' children, and not everyone.

## 8.11 The Case Against PDP: Scripture

There are Scriptural objections that can be made against PDP, however, and Talbott also pushes those. Talbott suggests Acts 17:28–29, Eph. 3:14–15 and 4:6 as Scriptural evidence for UPD (2013: 305).

For in him we live, and move, and have our being; as certain also of your own poets have said, For we are also his offspring. Forasmuch then as we are the offspring of God, [...] (Acts 17:28–29, NIV)

For this reason I bow my knees before the Father, from whom every family in heaven and on earth is named (Eph. 3:14–15)

[O]ne God and Father of all, who is over all and through all and in all. (Eph. 4:6)

We might also add:

Have we not all one Father? Has not one God created us? (Mal. 2:10)

I'm not persuaded that these texts do show that God is the father of all. The Greek for 'offspring; in Acts 17, *genos*', is, like *sperma* (seed), a less personal term than *teknon* (child) or *huios* (son) —its connotation is primarily biological in sense and so the description of the relation in Acts 17:28–29 lacks connotation of parental intimacy.<sup>62</sup> We would say that a spider has 'offspring', but it is strange to speak of a spider's 'children'.

Eph. 3:14–15 could simply refer to the patterning of all families after the fatherhood of God; and paying careful attention to the context of Eph. 4:6 makes it permissible to restrict the 'of all' to

---

<sup>62</sup> Rom. 9:7 is a good example of a place where this difference in connotation is exploited: 'neither because they are Abraham's seed are they all children [of Abraham]' (my translation).

mean ‘of all of us’—in v. 3 Paul’s concern is with the unity of Christian believers, and it is to this end that he emphasises the oneness of their calling in vv. 4–6.

Finally, Malachi 2:10 can be parsed as follows: ‘Have we, Israel, not all one Father? Has not one God created us *qua* nation-state?’. In short, the text is referring to the special relationship that Israel had to Yahweh, not that all mankind has to God. This national reading is supported by the fact that Malachi is attempting to extend the rebuking directed at the tribe of Levi in 1:6–2:9 to Judah and Israel more generally in 2:10–16.

That said, it is consistent with my position to grant that there are texts that speak of God’s being the father of more than merely the elect or believers. The Calvinist theologian John Murray (2009 [1955]) claims that there multiple senses in which we can refer to God as ‘Father’. Moreover, he even thinks that there is a sense in which God’s creating man *is* sufficient for God’s being the father of man. But he is keen to stress the distinction between the creative and adoptive senses of God as ‘Father’. Of the former sense he says,

Creatively and providentially he gives to all men life and breath and all things. [...] [I]t may be scriptural to speak of this relation which God sustains to all men in creation and providence as one of fatherhood and therefore of universal fatherhood. (2009 [1955]: 127–8).

But the adoptive sense describes ‘that most specific and intimate relationship which God constitutes with those who believe in Jesus’ name’ and Murray believes that to conflate this sense with the former one ‘means the degradation of this highest and richest of relationships to the level of that relationship which all men sustain to God by creation.’ (2009 [1955]: 128).

For my part, I am happy to say that there is a *sense* in which God is the father of the reprobate. This sense would be *analogous* to the true sense of fatherhood which God has only to the believers. Indeed, I can happily accept that there are numerous senses of fatherhood and that, relative to these different senses, the reprobate might often come out as God’s child. But it has been the presupposition of this chapter that *real* fatherhood commits one to loving one’s children equally, and therefore, whatever weaker or analogous senses there are in which all mankind are God’s children, because he does not love all mankind equally (as I have endeavoured to show), they are not all his children in the *real* sense.

## **8.12 God’s Love for the Reprobate**

This chapter is therefore a sound rebuttal to those like Channing, who, as we saw in ch. 3, insist that God's unconditional decrees of reprobation constitute a splendid failure of paternity. They would only be such a failure if God were the father of all, but we have good reason, quite independently of any Calvinist commitment, to be suspicious of the idea that he should be.

The material of this chapter also puts us in a position to comment on the question of God's love for the reprobate. As we saw, Frankfurt defined 'love' in terms of desire. To love someone is to desire to satisfy their interests. Can God therefore love the reprobate, even though he is not their Father? There is no reason why not. God can desire to satisfy the interests of the reprobate, and what is in the reprobate's best interest is his salvation, even though he has not decreed to satisfy that interest. A reason why he did not decree it was proposed in the previous chapter: the salvation of all would preclude certain goods to the elect. And, of course, the love God has for the reprobate will be non-paternal, for, being excluded from the family of God, the reprobate will therefore not be loved as much. But the love need be no less real on that account.

## Chapter 9—Conclusion

That brings this thesis to a close. In the course of the thesis I have (i) stated the best way of understanding theological determinism, (ii) grounded the doctrine in the Reformed tradition, and responded to the key objections made by critics: (iii) that determinism is inconsistent with moral responsibility, (iv) and that it undermines the goodness, love, and fatherhood of God.

In this conclusion, I merely note what I take to be the two most important areas for further research. The first is (v) the positive case for theological determinism. The posture of this thesis has been a purely defensive one. But there are many considerations that can be brought in favour of theological determinism: God's omnipotent power is weaker on Arminianism than on Calvinism, because God cannot directly bring about free decisions. This has ramifications for prayer: if God is restrained by free human choices, how much confidence can we place in him to answer prayers about such things? On Arminianism, one is saved by acting as an unmoved mover and choosing Christ. Doesn't this mean there is an aspect of the process of salvation for which God doesn't get glory, and doesn't that license proud boasting on man's part, in a way that the apostle Paul is keen to rule out? (Rom. 3:27) All of these arguments and more need to be brought to the table for a full assessment of the Calvinist position to be made.

The other important area that should be touched on is (vi) the Calvinist approach to the problem of evil. In this thesis, my focus was on the problem of *Hell*, not the problem of evil more generally. But a comprehensive Calvinist theodicy answering to both this-worldly and next-worldly evils appears a promising and necessary prospect. Much material may simply carry over: just as God's hatred of sin can be seen in the destruction of the reprobate, so it can also be seen in the natural disasters which humankind is subject to. And so on for all the other aspects of God's character that can be displayed in such things. Although there are differences one must be sensitive to. In the afterlife, the veil is lifted—the elect see God, and understand their relation to him, and can be counted on to infer correctly and prioritise righteously. Not so in this life: many spare little thought for the Lord God, and so may not perceive a divine display in some catastrophe. The theodicy must therefore accommodate imperfection in perception. Another big difference is that in the afterlife evils are distributed rigidly. No evil ever befalls the saved on the New Earth—it is all apportioned to the damned—and that permits the elect to be grateful and appreciative on account of the disparity. Yet evil befalls believer and unbeliever alike on this earth, and thus justifying goods that exploit the great contrast between the glorified and the damned can't be straightforwardly brought over to a this-worldly context.



## Bibliography

- Adams, Marilyn McCord. (1975) 'Hell and the God of Justice' *Religious Studies* 11, 433–47.
- . (1993) 'The Problem of Hell: A Problem of Evil for Christians' in *Reasoned Faith* (ed. Eleonore Stump) Ithaca: Cornell University Press.
- Adams, R. M. (1977) 'Middle Knowledge and the Problem of Evil' *American Philosophical Quarterly* 14(2), 109–117.
- Alighieri, Dante. (1867) *Divine Comedy* (trans. H. Longfellow) Many editions.
- Alston, William. (1991) 'The inductive argument from evil and the human cognitive condition.' *Philosophical Perspectives* 5, 29–67.
- Anderson, James and Paul Manata. (2017) 'Determined to Come Most Freely: Some Challenges for Libertarian Calvinism' *Journal of Reformed Theology* 11, 272–297.
- Anderson, James. (2016) 'Calvinism and the First Sin' in *Calvinism and the Problem of Evil* eds. D. Johnson & D. Alexander (Eugene: Pickwick Publications), 200–229.
- Anon. (2010) *The Canons of Dort*. Penascola: Chapel Library.
- Anscombe, G. E. M. (1957). *Intention*. Harvard University Press.
- Anselm. (1890) *Cur Deus Homo* London: Griffith, Farren, Okeden & Welsh.
- Arminius, Jacobus. (1853) *The Works of James Arminius, vol. 1*. Auburn and Buffalo: Derby, Miller and Orton.
- Asimov, Isaac. (1995) *I*, Asimov New York: Bantam.
- Astrain, A. (1908) 'Congregatio de Auxiliis' in *The Catholic Encyclopedia*. New York: Robert Appleton Company. (Retrieved September 13, 2019 from New Advent: <http://www.newadvent.org/cathen/04238a.htm>)
- Austin, John Langshaw. (1961) 'Ifs and cans' in his *Philosophical Papers*. Oxford: Oxford University Press.
- Baggett, David & Walls, Jerry. (2011) *Good God: The Theistic Foundations of Morality*. Oxford: Oxford University Press.
- Bailey, Andrew. (2012) 'Incompatibilism and the past' *Philosophy and Phenomenological Research* 85(2), 351–376.
- Ballor, J., M. Gaetano, and D. Sytsma. (forthcoming) *Beyond Dordt and De Auxiliis: The Dynamics of Protestant and Catholic Soteriology in the Sixteenth and Seventeenth Centuries*. Boston: Brill.
- Bawulski, Shawn. (2010) 'Annihilationism, Traditionalism, and the Problem of Hell' *Philosophia Christi* 12(1), 61–79.
- . (2013) 'Reconciliationism, A Better View of Hell: Reconciliationism and Eternal Punishment' *Journal of the Evangelical Theological Society* 56(1), 123–38.
- Berofsky, Bernard. (2006) 'Classical compatibilism: not dead yet' in *Moral Responsibility and Alternative Possibilities*, eds. David Widerker and Michael McKenna. Aldershot: Ashgate Publishing Limited.
- Bignon, Guillaume. (2018) *Excusing Sinners and Blaming God: A Calvinist Assessment of Determinism, Moral Responsibility, and Divine Involvement in Evil*. Eugene: Pickwick Publications.
- Blocher, Henri. (1992) 'Everlasting Punishment and Problem of Evil' in *Universalism and the Doctrine of Hell*, ed. Nigel M. de S. Cameron. Carlisle: Paternoster.
- Calvin, John. (1846 [1536]) *Institutes of the Christian Religion* trans. H. Beveridge. Edinburgh: Calvin Translation Society.
- . (1961 [1552]) *Concerning the Eternal Predestination of God* trans. J. K. S. Reid. London: James Clarke.

- . (2010 [1558]) *The Secret Providence of God* trans. K. Goad. Wheaton: Crossway.
- Carroll, Lewis (1899) *The Lewis Carroll Picture Book* (ed. S. D. Collingwood) London: T. Fischer Unwin.
- Chan, Godwell (1996) ‘Luther on Free Will’ *The Trinity Review*, Jan., 1–5.
- Channing, William. (1841) *The Works of William E. Channing, Vol. 1* (Boston: James Munroe and Company)
- Clark, Kelly James. (1995) “I Believe in God the Father, Almighty.” *International Philosophical Quarterly* 35(1), 59–69.
- Cottingham, John. (1986) “Partiality, Favouritism and Morality.” *The Philosophical Quarterly* 36, 357–373.
- Craig, William Lane. (1989) “‘No Other Name’: A Middle Knowledge Perspective on the Exclusivity of Salvation through Christ.” *Faith and Philosophy* 6(2), 172–188.
- . (2000) *The Only Wise God*. Eugene: Wipf and Stock.
- . (2016). *God Over All: Divine Aseity and the Challenge of Platonism*. Oxford: Oxford University Press.
- Crisp, Oliver. (2003) ‘Augustinian universalism’ *International Journal for Philosophy of Religion* 53 (3): 127–145.
- . (2010) ‘Is universalism a problem for particularists?’ *Scottish Journal of Theology* 63 (1): 1–23.
- Cunningham, William. (1862) *The Reformers; and the Theology of the Reformation*. Edinburgh: T & T Clark.
- Demetriou, Kristin. (2010) ‘The Soft-Line Solution to Pereboom’s Four-Case Argument’ *Australasian Journal of Philosophy* 88(4), 595-617.
- Dennett, Daniel. (1984) ‘I could not have done otherwise—so what?’ *Journal of Philosophy* 81(10), 553–565.
- Edwards, Jonathan (1977 [1754]) *The Works of Jonathan Edwards, Volume 1: Freedom of the Will* Yale: Yale University Press.
- . (1974a [1834]) *The Works of Jonathan Edwards, vol. 1*. Edinburgh: Banner of Truth (reprint).
- . (1974b [1834]) *The Works of Jonathan Edwards, vol. 2*. Edinburgh: Banner of Truth (reprint).
- Eef Dekker. (1996) ‘Was Arminius a Molinist?’ *The Sixteenth Century Journal*, 27(2), 337–352.
- Erasmus, Desiderius (1969 [1524]) ‘On the Freedom of the Will: A Diatribe or Discourse’ in *Luther and Erasmus: Free Will and Salvation* eds. A. N. Marlow & B. Drewery. London: SCM Press.
- Fesko, J. V. (2014) *The Theology of the Westminster Standards*. Wheaton: Crossway.
- Fischer, John Martin. (2004) ‘Responsibility and Manipulation’ *Journal of Ethics* 8, 145–77.
- . (2010) ‘The Frankfurt cases: The moral of the stories’ *Philosophical Review* 119(3), 315-336.
- . (2011) ‘The zygote argument remixed’ *Analysis* 71, 267–272.
- Fischer, John Martin & Ravizza, Mark. (1998). *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge University Press.
- Fisher, John Martin & Patrick Todd. (2015) ‘Introduction’ in *Freedom, Fatalism, and Foreknowledge*. Oxford: Oxford University Press, 1–38.
- Frankfurt, Harry. (1969) ‘Alternate possibilities and moral responsibility’ *Journal of Philosophy* 66(3), 829–839.
- . (2004) *The Reasons of Love*. Princeton: Princeton University Press.
- Franklin, Christopher Evan (2016) ‘If Anyone Should Be an Agent-Causalist, then Everyone Should Be an Agent-Causalist’ *Mind* 125(500), 1101-1131.

- Gatiss, Lee (2009) 'The Manifesto of the Reformation – Luther vs. Erasmus on Free Will' *The Churchman* (123:3) 203–225.
- Gill, John. (1855) *The Cause of God and Truth*. London: W. H. Collinridge.
- Goetz, Stewart. (2008) *Freedom, Teleology, and Evil*. London: Continuum.
- Gould, Paul M. (2014) (ed.) *Beyond the Control of God? Six Views on the Problem of God and Abstract Objects*. New York: Bloomsbury Publishing.
- Helm, Paul (2010) 'Editor's Introduction' in *The Secret Providence of God* by John Calvin. Wheaton: Crossway.
- Himma, Kenneth (2003) 'Eternally Incurable: the Continuing-sin Response to the Proportionality Problem of Hell' *Religious Studies* 39(1), 61–78.
- Hobbes, Thomas. (1651) *Leviathan*. Various editions.
- Hoefer, Carl. (2016) 'Causal Determinism' in *The Stanford Encyclopedia of Philosophy* at <<https://plato.stanford.edu/archives/spr2016/entries/determinism-causal/>>.
- Hume, David (1779) *Dialogues Concerning Natural Religion* (2<sup>nd</sup> ed.) London.
- Johnson, Daniel. (2016) 'Calvinism and the Problem of Evil: A Map of the Territory' in *Calvinism and the Problem of Evil* eds. David E. Alexander & Daniel M. Johnson. Eugene: Pickwick Publications, 19–55.
- Jordan, Jeff. (2012) 'The Topography of Divine Love.' *Faith and Philosophy* 29(1), 53–69.
- . (2015) 'The Topography of Divine Love: A Reply to Thomas Talbott' *Faith and Philosophy* 32(2), 182–187.
- Kane, Robert. (1998) *The Significance of Free Will*. Oxford: Oxford University Press.
- Kenny, Anthony. (1975) *Will, Freedom and Power*. Oxford: Basil Blackwell.
- Kearns, Stephen (2012) 'Aborting the zygote argument' *Philosophical Studies* 160(3), 379–389.
- Kershnar, Stephen (2018) *Total Collapse: The Case Against Responsibility and Morality*. Cham: Springer.
- King, M. (2013) 'The Problem with manipulation' *Ethics*, 124(1), 65–83.
- Kratzer, Angelika. (1977) 'What must and can must and can mean' *Linguistics and Philosophy* 1, 337–355.
- . (1981) 'The notional category of modality' in *Words, Worlds, and Contexts: New Approaches in Word Semantics*, eds. Hans. J. Eikmeyer and Hannes Rieser. Berlin: de Gruyter.
- . (2012a) *Modals and Conditionals*. Oxford: Oxford University Press.
- . (2012b) 'What must and can must and can mean' in *Modals and Conditionals*. Oxford: Oxford University Press.
- Kvanvig, Jonathan (1993) *The Problem of Hell* Oxford: Oxford University Press.
- Lehrer, Keith. (1968) 'Cans without ifs' *Analysis* 29(1), 29–32.
- Lewis, David. (2002) 'Evil for Freedom's Sake?' in his *Papers in Ethics and Social Philosophy*, vol. 3. Cambridge: Cambridge University Press.
- . (1976) 'The paradoxes of time travel' *American Philosophical Quarterly* 13(2), 145–52.
- Luther, Martin (1957 [1525]) *The Bondage of the Will* trans. J. I. Packer & O. R. Johnston. London: James Clarke.
- . (1857) *The Table Talk of Martin Luther*, trans. W. Hazlitt. London: H. G. Bohn.
- Lycan, William G. (1987) *Consciousness*. Cambridge: MIT Press.
- MacDonald, George (2008 [1858]) *Phantastes*. London: Paternoster.
- Maier, John. (2013) 'The Agentive Modalities' *Philosophy and Phenomenological Research* 90(1), 113–134.
- Matava, R. J. (2016) *Divine Causality and Human Choice: Domingo Banez, Physical Premotion and the Controversy De Auxiliis Revisited*. Boston: Brill.

- Matheson, Benjamin (2016). 'In defence of the Four-Case Argument' *Philosophical Studies* 173(7), 1963-1982.
- McCann, Hugh J. (1995) 'Divine Sovereignty and the Freedom of the Will' *Faith and Philosophy* 12(4), 582-598.
- . (1998) *The Works of Agency*. Ithaca: Cornell University Press.
- . (2005) 'The Author of Sin?' *Faith and Philosophy* 22(2), 144-159.
- McKenna, Michael. (2008) 'A Hard-Line Reply to Pereboom's 4-Case Manipulation Argument' *Philosophy and Phenomenological Research* 77, 142–59.
- . (2012) 'Moral Responsibility, Manipulation Arguments, and History: Assessing the Resilience of Nonhistorical Compatibilism' *Journal of Ethics* 16: 145–174.
- Mele, Alfred. (2006) *Free Will and Luck*. Oxford: Oxford University Press.
- . (2017) *Aspects of Agency*. Oxford: Oxford University Press.
- Molina, Luis de. (1988 [1588]) *On Divine Foreknowledge: Part IV of the Concordia*, trans. A. Freddoso. Ithaca: Cornell University Press.
- Murray, John. (2009 [1955]) *Redemption: Accomplished and Applied*. Edinburgh: Banner of Truth.
- Murray, Michael. (1999) 'Heaven and Hell' in his *Reason for the Hope Within* (ed.). Grand Rapids: Eerdmans.
- Nagel, Thomas. (1986) *The View from Nowhere*. Oxford: Oxford University Press.
- Nelkin, Dana (2019) 'Moral Luck' in *Stanford Encyclopedia of Philosophy* (at <https://plato.stanford.edu/archives/sum2019/entries/moral-luck/>)
- Nozick, Robert. (1974) *Anarchy, State, and Utopia*. New York: Basic Books.
- Parker, R. (2013) 'Deep and Wide: A Response to Jeff Jordan on Divine Love' *Faith and Philosophy* 30(4), 444–461.
- Pereboom, Derk (2001). *Living Without Free Will*. Cambridge: Cambridge University Press.
- . (2005) 'Free will, evil, and divine providence' in *God and the Ethics of Belief*, eds. Andrew Chignell & Andrew Dole. Cambridge: Cambridge University Press.
- . (2014). *Free Will, Agency, and Meaning in Life*. Oxford: Oxford University Press.
- Piper, John. (2000) *The Pleasures of God* (Rev. ed.). New York: Multnomah Books.
- Plantinga, Alvin. (1974) *The Nature of Necessity*. Oxford: Oxford University Press.
- . (1986) 'On Ockham's way out' *Faith and Philosophy* 3(3), 235–269.
- Queloz, Matthieu. (2017). 'Two Orders of Things: Wittgenstein on Reasons and Causes' *Philosophy* 92(3), 369-397.
- Raven, Michael J. (2015) 'Ground' *Philosophy Compass* 10(5), 322–333.
- Rea, Michael (2007) 'Metaphysics of Original Sin' in *Persons: Human and Divine* eds. Peter van Inwagen and Dean Zimmerman. Oxford: Oxford University Press.
- Rowe, William. (1979) 'The Problem of Evil and Some Varieties of Atheism' *American Philosophical Quarterly* 16(4), 335–341.
- . (1991) 'Ruminations about evil' *Philosophical Perspectives* 5, 69–88.
- . (1996) 'The evidential argument from evil: A second look' in *The Evidential Argument From Evil*, ed. Daniel Howard-Snyder. Bloomington: Indiana University Press.
- Rudy-Hiller, Fernando (2018) 'The Epistemic Condition for Moral Responsibility' in *The Stanford Encyclopedia of Philosophy* at <<https://plato.stanford.edu/archives/fall2018/entries/moral-responsibility-epistemic/>>.
- Saville, Andy. (2005) 'Hell without Sin: A Renewed View of a Disputed Doctrine' *Churchman* 119, 243–61.
- . (2007) 'Reconciliationism: A Forgotten Evangelical Doctrine of Hell' *Evangelical Quarterly* 79(1), 35–51.

- Schaff, Philip. (1931) *The Creeds of Christendom, Volume I* (6<sup>th</sup> ed.) USA: Harper & Brothers.
- . (1977) *The Creeds of Christendom, Volume III* (4<sup>th</sup> ed.) Grand Rapids: Baker.
- Sehon, Scott (2016) *Free Will and Action Explanation: A Non-Causal, Compatibilist Account*. Oxford: Oxford University Press.
- Seymour, Charles (1998) ‘Hell, Justice and Freedom’ *International Journal for the Philosophy of Religion* 43(2), 69–86.
- . (2000) ‘A Craigian Theodicy of Hell’ *Faith and Philosophy* 17(1): 103–115.
- Strawson, Peter. (1993 [1962]) ‘Freedom and Resentment’ in *Perspectives on Moral Responsibility* eds. Fischer and Ravizza. Cornell: Cornell University Press, 45–66.
- Swinburne, Richard. (1983) ‘A Theodicy of Heaven and Hell’ in *The Existence and Nature of God*, ed. Alfred J. Freddoso. Notre Dame: University of Notre Dame Press.
- Talbott, Thomas (2013) ‘The Topography of Divine Love: A Response to Jeff Jordan’ *Faith and Philosophy* 30(3), 302–316.
- Todd, Patrick. (2011) ‘A new approach to manipulation arguments’ *Philosophical Studies* 152(1), 127–133.
- . (2013) ‘Defending (a modified version of) the Zygote Argument.’ *Philosophical Studies* 164(1), 189–203.
- van Asselt, Willem, J. Martin Bac, & Roelf T. te Velde (2010) *Reformed Thought on Freedom: The Concept of Free Choice in Early Modern Reformed Theology*. Grand Rapids: Baker Academic.
- van Inwagen, Peter (1983) *An Essay on Free Will* Oxford: Oxford University Press.
- . (1984) ‘Dennett on ‘Could have done otherwise’’ *The Journal of Philosophy* 81(10), 565–567.
- . (1999) ‘Moral Responsibility, Determinism, and the Ability to do Otherwise’ *Journal of Ethics* 3, 341–350.
- . (2009) ‘God and other uncreated things’ in Kevin Timpe (ed.), *Metaphysics and God: Essays in Honor of Eleonore Stump*. London: Routledge.
- Vetter, Barbara. (2015) *Potentiality*. Oxford: Oxford University Press.
- Vihvelin, Kadri (2013) *Causes, Laws, & Free Will*. Oxford: Oxford University Press.
- Wainwright, William. (1988) ‘Original Sin’ in *Philosophy and the Christian Faith* ed. Thomas Morris. Notre Dame: University of Notre Dame Press.
- Wainwright, William. (2003) ‘Jonathan Edwards and the Doctrine of Hell’ in *Jonathan Edwards: Philosophical Theologian* eds. Paul Helm and Oliver Crisp. Aldershot: Ashgate.
- Werndl, Charlotte (2017) ‘Determinism’ in *The Routledge Companion to Free Will*. London: Routledge.
- Wesley, John. (1741) *Free Grace: A Sermon Preached at Bristol*. Boston: T. Fleet.
- . (1997) *The Complete Works of John Wesley, Volume 10: Letters, Essays, Dialogs and Addresses*. Albany: AGES Software.
- Widerker, David and McKenna, Michael. (2006) ‘Introduction’ in *Moral Responsibility and Alternative Possibilities*, eds. David Widerker and Michael Mckenna. Aldershot: Ashgate.
- Wykstra, Stephen. (1984) ‘The Humean obstacle to evidential arguments from suffering: On avoiding the evils of “appearance”’ *International Journal for Philosophy of Religion* 16(2), 73–93.