

# Healthy functioning as the key to fairness in a divinely determined world

DAVID B. HERSHENOV

*Department of Philosophy, Romanell Center for Clinical Ethics and the Philosophy of Medicine, 135 Park Hall, University at Buffalo, Buffalo, NY 14260, USA*  
e-mail: [Dh25@buffalo.edu](mailto:Dh25@buffalo.edu)

**Abstract:** Defending a traditional conception of Hell against the charge of unfairness is difficult for any theorist. It would seem to be even more difficult for the Christian semi-compatibilist who believes we can be responsible when determined. If God determined the wrongdoer's choices, then it seems unfair that he is punished for his actions that were manipulated rather than autonomous. I'll argue that healthy mental functioning can explain how some causally determined actions are manipulative while others are under our control. I'll also argue that God can punish even if He lacks the standing to blame.

## Introduction

It is difficult for any Christian to justify the fairness of an eternity in Hell. It would seem nearly impossible to justify such a fate if God determined the world and thus the character and conduct of those He damns. Nevertheless, I will argue that a semi-compatibilist – someone who thinks determinism doesn't undermine responsibility – can justify the fairness of divine punishments in a divinely determined world.<sup>1</sup> The key will be to draw upon ideas in the philosophy of medicine, a field that the philosophy of religion has rarely engaged. The idea of healthy mental functioning will be essential to making sense of how some causally determined actions are manipulative while others are expressive of our nature or essence and under our control.<sup>2</sup> We are accountable for the latter.

I'm not the first to appeal to the notion of healthy functioning to shed light upon autonomy. Although she doesn't consider the fairness of divine punishments for divinely caused actions, Sarah Buss (2012a) shows how one can be determined but nonetheless autonomous and accountable as long as one is functioning in a healthy manner. She has demonstrated how fertile the notion of health can be

for compatibilist, autonomy, and accountability research. I want to extend her basic insight about healthy functioning to the philosophy of religion's concern with the unfairness of Hell. To do so, I will have to reformulate Buss's account because she relies upon an unhelpful notion of the self. I'll modify her account, dropping her reliance upon a true or authentic self. Not only is the notion of 'self' woefully ambiguous but even the disambiguated senses don't seem necessarily tied to our essence and so aren't helpful in showing how healthy actions are expressive of our nature and mentally diseased actions are alien to us.

I will draw upon some recent work in the animalist personal identity literature to tie healthy functioning to our nature, thus enabling us to distinguish actions that are foreign to our nature from those that are not. We're essentially living beings with biological persistence conditions, while pathologies<sup>3</sup> are defined, in part, as involving subnormal contributions to survival (Boorse (1977); Wakefield (1992)). This renders pathologies foreign to us as they are by definition at odds with our continued persistence and instantiation of our nature. Healthy actions are expressive of our nature as they are designed to promote our persisting as living animals and consequently will be under our control in the sense relevant to autonomy and accountability. I will then extend my claims to animalism's opponents, advocates of psychological theories of personal identity. I will show how they too can connect healthy functioning to a person's nature and persistence conditions, which they understand in terms of psychology rather than biology. I will then be able to show how a range of theories render a health-based account of autonomy compatible with the theist's concerns about fairness.

My focus will mostly be upon two problems of unfairness in a divinely determined world. The first and main problem is that if God determined the character and choices of wrongdoers, then it seems unfair that they be punished at all for their actions weren't autonomous. Any punishment of the divinely manipulated would appear unwarranted. Even if divinely determined wrongdoers aren't manipulated, there is a closely related second charge of unfairness having to do with God lacking the standing to punish those whose conduct He causally determined (Todd (2012)).

Appealing to a health-based, proper functioning account of autonomy, I'll argue that the divinely determined creature is just as autonomous when the determinism is supernatural as when it is natural in its origins. I suspect philosophers don't see this because they're misled by God's intentions being loosely analogous to genuine cases of evil scientist-like manipulation where the manipulator's intentions replace those that the manipulated would have had. But it is a mistake to believe that supernatural determinism differs in any relevant respect from the threat to our autonomy brought by nature's mindless determinism. Once natural determinism is recognized as compatible with autonomy and accountability, divine determinism can be likewise accommodated.

The second charge that it is unfair for God to punish those He caused to engage in misconduct will be met by the claim that posthumous punishment is a necessity

and God is the only one who can provide it. God's alleged complicity with the wrongdoers in a causally determined world leads philosophers to think wrongly that it undermines His standing to both blame and punish. But standing to blame and standing to punish can come apart. What's important is that God punish those He caused to do wrong when He is the only one that can. It's no more wrong for God to do so than it is for parents to punish an adult child whose bad character they earlier formed when they find themselves later being the only ones so situated to make their child pay his debt and undergo what might be a painful process of reform, atonement, and reconciliation.<sup>4</sup>

### **Causally determined autonomy**

Buss provides the theoretical framework for our solution to the problem of fairness in a divinely determined world. She recognizes, unlike some free will source sceptics, that accountability demands self-determination, not *determination of the self*. So the self can be externally formed and causally determined, even intentionally determined, without our accountability and autonomy being excluded.

Buss acknowledges that there are many conceptions of autonomy and accountability. She is interested in what makes someone self-governing and accountable that is more than acting voluntarily, intentionally, and for a reason. Buss's thesis is that autonomy is grounded in the healthy functioning that passively causes our intentions and actions rather than to be found in appeals to idealized conceptions of agency. Where I will eventually differ with Buss is that she advocates a 'true self' while I find such talk to be misleading, terribly ambiguous, and trivial or false when disambiguated and explicated. A typical claim of hers is that 'We are more *truly ourselves* when we are not afflicted with an illness or disability and it is this normative aspect of our identity – our identity as a representative member of our species – that is the key to the special sense in which autonomous agents determine their own actions' (Buss (2012a), 650). My aim is to replace Buss's unhelpful talk about a true self with a helpful account based upon our nature as animals that renders diseases foreign and actions based upon healthy functioning as autonomous as they emanate from our nature.

I am, however, convinced by Buss's rejection of the thesis that exercising autonomy and being accountable involves some sort of idealized deliberation or endorsement. She argues first that there could be accountability without reflection and deliberation. One is accountable even when one acts in a habitual, spontaneous, or careless manner. She also rejects understanding the exercise of autonomy to consist in higher-order attitudes (desires or values) endorsing lower-order preferences or desires. She insists that autonomy can be found in the absence of endorsement as autonomous agency is compatible with ambivalence, regret, disappointment, frustration, and self-criticism. Akratics are accountable and blameworthy, even if self-repudiating. So I can be blamed for taking someone else's

piece of chocolate cake when in a moment of weakness I give in to a temptation while knowing full well that I shouldn't. Thus neither endorsement nor deliberation is a necessary condition for autonomy. Nor are they sufficient as deliberation and endorsement can be manipulated. Buss points all of this out before even helping herself to the well-known regress problem to show the failure of idealized agency accounts of autonomy (*ibid.*, 656).

Buss maintains that the key to autonomy lies in the healthy passive unchosen background that influences agents' choices and actions. The agent is autonomous when she can be 'identified' with the direct, purely causal, non-rational influences on the formation of her intentions. Even out of character actions, as well as characteristic ones, are autonomous when caused by healthy psychological-physiological functioning that will provide the basis for minimal flourishing for species members.

Buss observes that mental disease is like a 'hostile take-over' (*ibid.*, 668). She illustrates this through descriptions of obsessive-compulsive pathologies and eating disorders. Such mentally diseased people can't help but think that they must constantly wash germs off their hands or fast as 'punishment' for being overweight. Autonomous agents can take certain diseases into account while mental diseases determine what agents take under consideration. Owing to failures of their background cognitive and affective faculties to function properly, the mentally diseased see certain things as reasons that the healthy do not. Buss points out that what counts as a consideration for acting due to a passively experienced background can colour the agents' perceptions and evaluations and determine how much weight is assigned to such reasons. Mental illnesses undermine our autonomy even when we carefully review our actions and endorse or identify with them. She claims that diseases are foreign to us as they result in actions not under our control or expressive of our true selves.

### **Health, autonomy, and our nature**

There is much that is attractive about Buss's account. Of course, as with all new theories, there will be problems that have to be worked out. I will focus upon how unhelpful the notion of the *self* is. Buss writes: 'In what follows I will try to show that it is this distinction between self and other determined behavior that is key to the distinction between autonomous and non-autonomous behavior' (*ibid.*, 660); 'She is not really *herself* when sick' (*ibid.*, 662); 'we are more truly *ourselves* when we are not afflicted with an illness or disability' (*ibid.*, 667). Unfortunately, Buss never provides an informative account of how to divide the person into true and foreign components, explaining why diseases are foreign and actions arising from them are not expressive of our selves. She says nothing about how we can appeal to our nature to justify such distinctions. She doesn't tie the self to anything essential. Buss does say that she means by self-governing what Jean Hampton does in the latter's discussion of 'selfless' behaviour that is

not expressive of the 'real and satisfactorily operating person' (*ibid.*, 667, n. 42). Alas, the reference to the passage in Hampton is unhelpful as the latter writes:

Just as a sculptor created a form out of a slab of rock, so too do people (in concert with the environment and their biology) *create a distinctive* way of interacting with, thinking about, and reacting to the world. It is this distinctiveness which each individual plays a major role in creating, that I call the 'self.' Whereas we say that we respect one another as 'persons', we say that we love or hate, approve or condemn, appreciate or dislike others' selves. (Hampton (1993), 150)

Hampton leaves me wondering who's doing the distinctive creating if it is not the self – is it the person or the human animal or some other individual?<sup>5</sup> How can people be different from selves? Why would we be more truly a self than a person? Moreover, if we create the self, then why is it not something contingent rather than deep and essential?

I am also sympathetic to Eric Olson's claims that talk of the 'self' is terribly ambiguous. He observes: 'Depending upon whom you believe, selves may be concrete, or abstract, material or immaterial, permanent or ephemeral, naturally occurring or human constructions, essentially subjective or publicly observable, the same or not the same things as people' (Olson (1998), 645). Even if we disambiguate 'self' talk, I don't see how any of the conceptions make some actions alien to us.

I fear that Buss's talk of a 'true self' and 'certain traits being more fully human' (Buss (2012a), 668) collapses into just the absence of malfunctions that are diseases. Diseases can't be contrary to our identity for we are certainly in possession of diseased properties. Christopher Boorse, the leading philosopher of medicine, believes that we are diseased if we have any diseased parts and none of us is completely without diseased parts (Boorse (1997), *Idem* (2014)). The diseased parts truly belong to us, i.e. they are not foreign parts as was the bullet in the leg of Sir Arthur Conan Doyle's Dr Watson. We can be irreversibly diseased, we can identify with our disease, we can be congenitally diseased, and we can manifest diseases from conception. Moreover, autonomy is compatible with out of character behaviour so that can't be the basis for diseases being foreign. Buss even admits that there can be diseases which are part of one's narrative identity and thus seem authentic.<sup>6</sup>

A 'core self' and 'true self' that is distinguishable from one's diseased parts or properties is either false talk or unhelpful metaphors. Since I accept an animalist account of personal identity, when I refer to *myself* I am not referring to any mental core but just the animal that I am. I can refer to my mindless embryonic self and to my future demented or comatose self. No 'core self' or 'true self' is picked out by first-person referential terms and they are the only unproblematic and unavoidable uses of 'self'.<sup>7</sup>

If there is no true or authentic or real self, how can we preserve the idea that disease-caused actions are alien, heteronomous, not under our control, and not

expressive of ourselves? I think an alternative account in the spirit of Buss's formulation can be provided. First, we must make sense of the alien nature of disease. Diseases are contingent traits so we must separate different types of an individual's contingent traits. The key to finding what in our nature makes diseases foreign is to appreciate our biological essence and persistence conditions: We are animals and are essentially alive.<sup>8</sup> It is the nature of the living to function in a certain way to remain in existence as living organisms. This is true regardless of species kind. Healthy cognitive and affective functions are defined, in part, in terms of species typical contributions (Boorse (1997)) or naturally selected contributions (Wakefield (2005)) to the organism's survival.<sup>9</sup> Diseases are in a sense at odds with our essence, nature, and persistence conditions even though we can possess them in ways we cannot possess other properties that are at odds with our nature such as being abstract, atemporal, or non-extended. Since diseases are typically defined in terms of suboptimal contributions to organism survival (and reproduction), this renders them opposed to our essence in ways that other contingencies like different hair colours aren't. This allows us to see diseases as foreign for they undermine what is essential to us – life processes. Actions that are symptomatic of mental diseases are thus alien to us as they are in opposition by their nature to our persistence.

Not all theories of personal identity claim that we are essentially animals. Some even maintain that we don't die *biological* deaths and thus there is a need for two senses of 'death' and different criteria of death.<sup>10</sup> I, on the contrary, hold that any theory of personal identity that has us possessing a living body should accept that we die biological deaths.<sup>11</sup> Death involves ceasing to instantiate life processes. So even if you believe that our psychology is essential to us and our biology isn't, you should accept that we die when our minds are extinguished. Since we cease then to exist, we cease then to instantiate life processes. The latter isn't rendered false by the possibility that what was our body could still exist and be alive after we cease to exist and die. To see this, consider the analogy that if your finger is bruised that means both you and your finger are bruised. Two distinct but overlapping entities can instantiate the same property – being bruised. If all of your body but your finger is destroyed, it would be correct to say that you were destroyed, not that you have become reduced to the size of your finger. So you cease to be bruised when you cease to exist, even though your bruised finger remains. Likewise, advocates of a psychological account of personal identity should claim that you cease to instantiate life processes and thus die when you cease to exist, even though your body may remain alive. So it is a mistake of the advocate of psychological theories of personal identity to claim that we need two definitions and two criteria of death – one for animals and one for persons. All that is needed are two criteria for going out of existence – one for those creatures that are essentially alive and contingently thinking and another for those that are essentially thinkers and only contingently living animals.

Still, it may be replied that even if we recognize that we can die despite not being identical to our animal, that isn't enough to tie healthy functioning to our nature and thus render unhealthy mental acts as foreign. But it is easy to adapt what I said about death being determined by suboptimal contributions to the human animal's survival to theories that deny we are identical to an animal. Disease can be considered as what undermines the essentially thinking person's survival. The same mental diseases that make it more likely that the human animal will die also make it more likely that the human person dies; and the human person will nearly always cease to exist when she dies. It is only in the hypothetical thought experiments that human animals cease to exist while human persons do not. The three most famous personal identity thought experiments – the transplant of your cerebrum and the destruction of your previous body, sustaining your (upper) brain in a vat after your body is destroyed, and the replacement of your organic body with inorganic parts – are taken to show that persons could survive the loss of their living animal body.<sup>12</sup> But every actual case of destroying the animal body destroys the material person.<sup>13</sup> To handle the merely possible cases, the pathology of a *person* can be reconstrued in terms of making certain suboptimal contributions to a person's survival rather than to sustain their biological life. This will not change psychiatry's bible – the Diagnostic and Statistical Manual of Mental Disorders (DSM) – nor make any of what we now recognize as unhealthy autonomy-undermining actions into healthy autonomy-preserving actions. So even if we accept a psychological account of personal identity, we can still understand some disordered actions as contrary to our nature as they make it less likely that we will continue to persist and instantiate our nature as essentially thinking beings.<sup>14</sup>

### **Determinism, manipulation, and punishment**

The Old Testament mentions God's hardening of Pharaoh's heart.<sup>15</sup> Contrast these interventions with God setting causal processes into motion eons earlier when the world was created that produced Pharaoh's bad character. It is this mean-spiritedness which leads the Egyptian ruler initially to mistreat the Israelites. It is later that God hardens Pharaoh's heart after Pharaoh has relented and decides to free the Hebrews. A Buss-inspired account can explain why Pharaoh is accountable for some of his actions but not others. The cases in which God hardened Pharaoh's heart would be manipulation. He shouldn't punish the manipulated.<sup>16</sup> They are to be excused. While Pharaoh may not be autonomous when God intervenes to harden his heart, Pharaoh could be held accountable for what he has earlier done as a result of God initially setting into motion events at the world's creation that lead Pharaoh initially to brutalize the Israelites. The reason that Pharaoh is accountable for his earlier misconduct is that his cognitive and affective faculties were then working as they were designed. It is only the interventionist God that later manipulated Pharaoh, hardening his

heart to prevent the exodus when he otherwise would have let the Jews go. The crucial difference between the manipulated and the autonomous in a deterministic world is that the former are not functioning in the cognitive and affective manner for which they were designed. The proper functioning of their faculties have been bypassed or overridden by interventions of the divine creator or evil neuroscientists.<sup>17</sup>

A healthy functioning-based account of autonomy and accountability also allows us to develop a response to Patrick Todd's doubts about divine punishments when the sins are divinely determined. Todd (2012) has cleverly extended manipulation arguments against compatibilists to God's standing to blame those who do wrong in a world that is divinely determined. It seems counterintuitive to blame someone for something one caused them to do. So if God 'has long ago . . . irrevocably determined everything that happens, including our bad actions' (*ibid.*, 2) and God stands ready to judge those He determined to do what they did, it intuitively seems unfair for them to be blamed. Even if someone else could blame them, the worry is that God doesn't have the moral standing to do so. Although Todd's article focuses on the 'unhappy tradition' (*ibid.*) wherein God blames us for what He determined to come about, he points out the same arguments can be extended to punishment. He explains:

[I]t would seem that the points that I have made concerning the possibility of God's blaming us could (*mutatis mutandis*) be applied to show the permissibility of God's *punishing us*. We could thus make the scenario more unsettling by imagining that God says 'What you have done is a horrible thing to have done, and its horribleness deserves punishment: 15 years in jail.' Even if you are convinced that (abstractly considered) 15 years in jail is a perfectly just punishment for such an act, wouldn't you think – quite rightly – that the whole game is rigged, and that the universe is deeply unfair? That the moral universe could be this way strikes me as intolerable, though I'm sure some could find it in their hearts to tolerate this. (*ibid.*, 16)

Todd could have made his case even stronger with an eternity in Hell rather than 15 years in prison awaiting those whom God Himself determined to do wrong. So as a semi-compatibilist, I admittedly have my work cut out for it is an uphill climb.<sup>18</sup> How I propose to convince readers that God's punishing those he caused to sin is not as dire a scenario as Todd imagines is to suggest that people are misled by these manipulation stories. Insofar as semi-compatibilists can accept natural determinism and maintain that people are responsible for their determined acts, they should deny that autonomy-undermining manipulation occurs because God has set events in motion at the time of Creation. Tell your favourite story about responsibility not being undermined by nature's causally determining you to do what you did and there is then no reason to change your judgement when the determinism is intentional in origin.

If we react differently to divine determinism than to natural (mindless) determinism, judging only the first to undermine the agent's accountability, we are inconsistent. Something must give. I recommend that we should only trust our judgement in the latter. Perhaps divine manipulation cannot teach us much



about the responsibility of determined agents because it introduces too many elements (complicity, hypocrisy, double effect, theodicy, problem of evil, retribution) that complicate the matter, keeping us from realizing that determinism by another agent is not autonomy undermining. I suspect that the main reason we should not trust such initial reactions to intentional determination scenarios is that they distort our judgement by annexing determinism to the instrumental use of another. If someone discovers or imagines that he has been determined to think what he does by another agent then he will (i) fear that he is being used by another and (ii) doubt that the determined intentions are really his own. These considerations will lead him to deny here that his intentions are autonomous while the same intentions produced blindly in a deterministic world would not. Doubts about 'ownership' of our intentions and fears about being instrumentally used by others lead to worries about our accountability being undermined by a powerful agent that are absent in natural, ultimately agentless determinism. My suspicion is that the different response to the intentional determinism is due in part to our thinking that we would have had different thoughts and have done something else if there hadn't ultimately been an external agent behind a causal chain operating upon us. But only in an indeterministic world with 'incompatibilist manipulation' could we have thought something differently in the absence of someone else causally determining our intentions. The compatibilist should thus treat their intentions in the same way regardless of whether the world is determined by blind nature or a divinity.

It might help us see the determining *agent's* actions as benign rather than manipulative if we imagine the overdetermination of our intentions. Conceive of what the alleged 'manipulating' agent does as mere overdetermination of what would occur in the absence of that person's causal involvement with your own intentions. First, fill in all the details of a semi-compatibilist story about your intentions emerging in the right way under your control in a world that is determined by blind nature. Then add a powerful person merely overdetermining an intention you would have had in the absence of this agent's intervention. I don't think you will see your autonomy threatened here because there is an agent rather than just mindless nature determining your actions.

Readers may agree with me that the divinely determined are autonomous and accountable but they may still share Todd's worries that God doesn't have the standing to blame and punish. I suspect that even if God can't blame those He causes to sin,<sup>19</sup> it is still true that He can punish them. My contention is that the correct approach is to realize that punishment involves not retribution but reform, atonement, and reconciliation. It is still punishment because this involves some intended harm (Boonin) in response to wrong.<sup>20</sup> Even if God is responsible for giving someone their bad character and can't blame that person because of His involvement, He can punish just as you can if you are to blame, in part, for the wrongdoing of your child. Imagine scenarios where your child cannot be turned over to the authorities and only you can punish him (reform him, make him

pay his debt, urge him to atone, paternalistically bring about his being harmed so he comes to know better how his victim felt, etc.). You can do that even if you are involved in his wrongdoing. It may be that you intentionally gave him a bad character to bring about some good or you (unlike God) aimed to have your child commit some evil as an end but have since reformed, or your child's wrongful act was an unintentional but foreseeable consequence of your rearing practices. The key is that you are the only one so situated to punish. God, of course, will be the only one positioned to punish someone in the afterlife.

A different objection against God having standing to punish is that He was neither the wronged party nor their agent.<sup>21</sup> A quick answer is that the human victims of the wrongdoer exist posthumously and would typically want the wrongdoer's apology, respect, and reconciliation, and would authorize or recognize God as the agent of this process. Second, God stands in an intimate relationship to his creations that makes the same acts that wrong them be wrongs to Him. Kvanvig offers as an analogy the intimacy, immediacy, and connectedness of parents to their newborn. God's relationship to us is even closer than that of the parents to their newborn, whom they created and sustained. A wrong to that child is arguably a wrong to the parents, at least when that child is young and being nurtured by them. God not only created everything in the universe but, according to the doctrine of divine conservation, sustains all of it at every moment of its existence (Kvanvig (1993), 36–37).

My conception of Hell has it serving not to prevent future mischief in the afterlife (Seymour (1997)) but as the location for the never-completed punishment for the original sin(s) on earth. Punishment in Hell need not involve unbearable suffering for that may not be compatible with God's loving character. But His loving nature is not separate from His just nature, so punishment is not neglected. However, the gap between justice and love and mercy isn't the wide one commonly supposed since God seeks a loving reconciliation with the unjust and mercy is always an option he can show the contrite.<sup>22</sup> Our psychology being what it is, full contrition and reconciliation may not come until long after death, hence the need for Purgatory.<sup>23</sup> God's foreknowledge of his determined creatures allows him to mercifully bestow upon his children as many chances to atone as are required<sup>24</sup> in Purgatory, thus avoiding the need for an escape from Hell as argued (Buckareff & Plug (2005)) or an abandonment of the doctrine of a Final Judgement (Seymour (1998)).

God's loving and merciful nature is such that He wants to spend eternity with us despite our sinful pasts. But some of us don't want to spend eternity with Him on His conditions. Some wrongdoers don't show remorse. They don't apologize and they don't seek reconciliation. It doesn't seem to be morally troubling to fail to forgive those who don't ask for forgiveness and show no contrition. Parole boards may keep two criminals in jail for different times despite their having conspired to commit the same crime. One is repentant and the other is not. The former goes free because he is apologetic, willing to make amends, and seeks to

be forgiven and reconcile with those he wronged. The other stays in jail. There is a parallel with the diabolically defiant in the afterlife. Their punishment cannot be objected to as excessive or unfair (Kershnar (2005, 2010); Lewis, (2007); Adams (1975)) as involving infinite punishments for finite harms. We're imagining wrongdoers forever refusing to repent and ask forgiveness. An eternity in Hell isn't then excessive for it is a response to an eternal rejection.<sup>25</sup>

God loves and respects the creatures that reject him which thus limits what He can do to enable them to avoid eternal punishment. So He won't override their autonomy. Thus semi-compatibilists have their own free will response to the problem of evil, and the problem of Hell is a version of that problem (Kvanvig (1993), 4). God won't manipulate His children into being good. Thus God respects their autonomy by not forcing some of his creations to join Him in Heaven. This is compatible with his loving nature, satisfying what Kvanig calls an issuant conception of Hell (*ibid.*, 112).

Is it unfair that God would make such obstinate and imprudent creatures rather than actualize a world where they would have different more 'Heaven-friendly' characters? Why should there even be a need for Hell or even Purgatory? David Lewis asks something similar when discussing limited punishment (which are punishments without the permanence of Hell and thus akin to those of Purgatory).

Presumably there is some great end that God has designed his creation to achieve, an end that is furthered by the repentance of those who failed the earthly test. An obvious rejoinder, from those of us who find no great value in incompatibilist freedom, is that God could have saved himself the trouble of limited punishment by setting up the causal conditions so that the resisters didn't go astray to begin with. (Lewis (2007), 235)

That is a good objection, one no doubt shared by many readers.<sup>26</sup> Perhaps a good place to start to begin to answer the question is Peter van Inwagen's (1995) discussion of earthquakes and floods. Maybe God couldn't have made a world that sustains us without them. Something similar may be true about obstinate personalities that reject God. Van Inwagen pointed out that it is hard to imagine that evolution could have occurred without thousands of years of suffering (van Inwagen (1995), 81). Perhaps we couldn't have evolved the psychological capacities we do have without some people coming to be disposed to defy God. Van Inwagen cautions that we who have never made a world shouldn't be so confident that we know what is possible and what is not (*ibid.*, 106).

Centuries before van Inwagen, Irenaeus even conjectured that necessarily it was the case that 'created things must be inferior to Him who created them . . . but insomuch as they are not uncreated, for this very reason they come short of perfection' (Irenaeus, *Against Heresies*, bk 4, ch. 38, p. 42; cited in Seymour (1997), 262). The resulting soul-making may involve scenarios where temptations and vices are real possibilities. So it is plausible that God couldn't have bestowed upon us the moral psychology required to get to Heaven without some of us becoming diabolically obstinate.<sup>27</sup>

Our Buss-inspired account means that even if the damned did complain about having a Heaven-excluding obstinate character, they couldn't protest that their autonomy was violated and thus they were being used as a mere means by God. If people in Hell identify with their character they may protest against their fate but not against the character that brought them there. It is perhaps more likely that others will claim they have been treated unfairly in the creation of their psychology. But it strikes me as dubious to claim that there is unfairness in character formation when it is not recognized by those who are allegedly unfairly treated and yet are well informed and would reject alternative characters. And if the logically possible alternatives where absolutely no one has Heaven-resistant characters are not metaphysically real options for reasons like those suggested by van Inwagen and Irenaeus, then the charge of unfairness is even less credible.

### Conclusion

Defending a traditional view of Hell is difficult enough. To do so from a semi-compatibilist position makes an uphill climb seem like virtually a vertical ascent. I hope to have shown that the philosophy of medicine provides a more informative guide than philosophers of religion were previously aware. Still, like most authors, I probably believe that I have climbed to heights greater than all of my readers think I have reached. It is far more likely that we all agree on what mountain I was trying to climb. I assumed the traditional view that Hell is occupied, inescapable, unpleasant, and that there is a final judgement at which our fate is sealed. But God's foreknowledge in a determined world enabled him to give his children years of posthumous opportunities needed for contrition in Purgatory. So God's unlimited mercy and forgiveness is compatible with people being sent to Hell on Judgement Day and not leaving. We are responsible for our fate even though God causally determined the outcome. We are no less responsible if the determinism is due to a divine agent rather than blind nature when it operates through our functioning in a healthy manner. It is not unfair that those who autonomously reject God are forever estranged from him in Hell. As C. S. Lewis memorably said: 'The gates of Hell are locked from the inside' (Lewis (2001), 127).

### References

- ADAMS, MARILYN MCCORD (1975) 'Hell and the God of Justice', *Religious Studies*, 11, 433–447.
- BAKER, LYNNE (2005) 'When do persons begin and end', *Distinguished Faculty Lecture*, <<http://people.umass.edu/lrb/files/bako5whebM.pdf>>.
- BOONIN, DAVID (2008) *The Problem of Punishment* (Cambridge: Cambridge University Press).
- BOORSE, CHRISTOPHER (1977) 'Health as a theoretical concept', *Philosophy of Science*, 44, 542–573.
- BOORSE, CHRISTOPHER (1997) 'A rebuttal on health', in J. Humber & R. Almeder (eds) *What is Disease?* (New York: Humana Press), 1–134.
- BOORSE, CHRISTOPHER (2014) 'A second rebuttal on health', *Journal of Medicine and Philosophy*, 39, 683–724.

- BUCKAREFF, ANDREI & PLUG, ALLEN (2005) 'Escaping hell: divine motivation and the problem of hell', *Religious Studies*, **41**, 39–54.
- BUSS, SARAH (2012a) 'Autonomous action: self-determination in the passive mode', *Ethics*, **122**, 647–691.
- BUSS, SARAH (2012b) *Ethics at Pea Soup Blog*, <<http://peasoup.typepad.com/peasoup/2012/10/ethics-at-pea-soup-sarah-buss-autonomous-action-self-determination-in-the-passive-mode-with-comment.html#more>>.
- FISCHER, JOHN MARTIN (2006) *My Way: Essays on Moral Responsibility* (Oxford: Oxford University Press).
- HAMPTON, JEAN (1993) 'Selflessness and the loss of self', *Social Philosophy and Policy*, **10**, 135–165.
- HERSHENOV, DAVID (2005) 'Do dead bodies pose a problem for the biological account of identity?' *Mind*, **114**, 31–59.
- HERSHENOV, DAVID (2010) 'Restitution and punishment', in J. Ryberg & A. Corlett (eds) *Punishment and Ethics: New Perspectives* (New York: Palgrave Macmillan Press), 33–51.
- IRENAEUS. *Against Heresies: The Ante-Nicean Fathers. Vol. 4.*
- KERSHNER, STEPHEN (2005) 'The injustice of hell', *International Journal for Philosophy of Religion*, **58**, 103–123.
- KERSHNER, STEPHEN (2010) 'Hell and punishment', in J. Buenting (ed.) *The Problem of Hell: A Philosophical Anthology* (Burlington VT: Ashgate), 115–132.
- KVANVIG, JONATHAN (1993) *The Problem of Hell* (Oxford: Oxford University Press).
- LEWIS, C. S. (2001) *The Problem of Pain* (New York: Harper Collins).
- LEWIS, DAVID (2007) 'Divine evil', in L. Anthony (ed.) *Philosophers without God: Meditations on Atheism and the Secular Life* (Oxford: Oxford University Press), 231–242.
- LIZZA, JOHN (2012) 'Brain death', in F. Gifford (ed.) *Handbook of the Philosophy of Science* (North Holland: Elsevier), 453–488.
- MCMAHAN, JEFF (2002) *The Ethics of Killing: Problems at the Margins of Life* (Oxford: Oxford University Press).
- OLSON, ERIC (1998) 'there is no problem of the Self', *Journal of Consciousness Studies*, **5**, 645–657.
- OLSON, ERIC (2013) 'The Person and the Corpse', in I. B. Bradley et al. (eds) *The Oxford Handbook of Philosophy of Death* (Oxford: Oxford University Press), 80–96.
- OLSON, ERIC (2015) 'Life after death and the devastation of the grave', in M. Martin & K. Augustine (eds) *The Myth of an Afterlife* (Lanham MD: Rowman & Littlefield), 409–423.
- PEREBOOM, DEREK (2001) *Living Without Free Will* (Cambridge: Cambridge University Press).
- SEYMOUR, CHARLES (1997) 'On choosing hell', *Religious Studies*, **33**, 249–266.
- SEYMOUR, CHARLES (1998) 'Hell, justice, and freedom', *International Journal for Philosophy of Religion*, **43**, 69–86.
- SIDER, TED (2002) 'Hell and vagueness', *Faith and Philosophy*, **19**, 58–68.
- TODD, PATRICK (2012) 'Manipulation and moral standing: an argument for incompatibilism', *Philosophers' Imprint*, **12**, 1–18.
- VAN INWAGEN, PETER (1995) *God, Knowledge and Mystery: Essays in Philosophical Theology*. (Ithaca NY: Cornell University Press).
- VEATCH, ROBERT (2005) 'The death of whole-brain death: the plague of the disaggregators, somaticists, and mentalists', *Journal of Medicine and Philosophy*, **30**, 353–378.
- WAKEFIELD, JEROME (1992) 'Disorder as harmful dysfunction: on the boundary between biological facts and social values', *American Psychologist*, **47**, 373–388.
- WAKEFIELD, JEROME (2005) 'Biological function and dysfunction', in D. Buss (ed.) *Handbook of Evolutionary Psychology* (New York: Oxford University Press), 878–902.
- WALLS, JERRY (2011) *Purgatory: The Logic of Total Transformation* (Oxford: Oxford University Press).

## Notes

1. Fischer defends the distinction between compatibilism and semi-compatibilism (Fischer (2006), 76–78).
2. I am using 'nature', and 'essence' as synonyms.
3. I am using 'pathology', 'disorder', and 'disease' as synonyms.
4. I'll be assuming more than defending here that punishment should involve a debt payment, atonement, and reconciliation. Hershenov offers defence more than an assumption of that approach (Hershenov (2010)). The basic idea is that the wrongdoer has a debt that must be paid to his victim(s) to restore their relationship with the wronged. Doing so willingly can be evidence of one's contrition and reform and warrant the restoration of the relationship sundered by the wrong.

5. And do we lack a self if it is not distinctive but has a duplicate? Surely not.
6. Buss realizes that we 'need to distinguish conditions of accountability from what we may call conditions of authenticity' for someone 'would be alienated from herself to want to be cured of what is so central to her way of being in the world' (Buss (2012b), 6).
7. Olson claims all the other uses of 'self' can be replaced with a more clear paraphrase.
8. The 'Termination Thesis' is that we go out of existence at death since we are essentially alive (Hershenov (2005); Olson (2013), *Idem* (2015)).
9. I write 'in part' because philosophers of medicine stress contributions to reproduction as well as survival. Since I don't think contributing to reproduction or even being able to evolve is essential to being an organism, I'm not construing reproductive pathologies to be at odds with our nature.
10. Defenders of psychological approaches to personal identity deny that we are identical to a human animal and so claim that there is a need for one conception and criterion of death that applies to us and another that applies to animals (Baker (2005); McMahan (2002); Veatch (2005); Lizza (2012)).
11. So psychological continuity accounts of personal identity, constitution accounts, as well as hylomorphic accounts, and even dualist theories that claim we are composed of a soul and body rather than are just identical to our soul, should accept that we die when our bodies die.
12. The cerebrum or upper brain is not a living human animal when it is in transit from one body to another. And replacing your organic parts will mean the death of your animal even if a psychologically functionally equivalent inorganic body guarantees the survival of your person.
13. Even Thomistic hylomorphism of the corruptionist version has human beings or persons going out of existence at death though a proper part of them, their soul, continues. It is only with resurrection that the deceased come back into existence.
14. There may be a religious notion of healthy or proper functioning that involves people being designed to reconcile with God which renders dysfunctional anyone who resists God. This, obviously, is not the sense of healthy functioning that will be relevant to autonomy in a fallen world. Those who eternally resist God may be imprudent but not mentally ill. By analogy, I know that I could have had better children who make me happier but I don't want God to roll back time to where I have different and better children. That may be imprudent, but given my character and love for my actual children, I don't want to live differently, even if I would be happier doing so. I am tempted to say that is imprudent but not autonomy undermining mental illness.
15. The relevant biblical passages are Exodus 4:21, 7:3, 9:12, 10:1, 19:21, 11:10, 14:8.
16. Maimonides notes that Pharaoh already deserved punishment before his heart was hardened.
17. A Buss-inspired, healthy functioning-based account of autonomy (Buss (2012a), 688-689) can also provide a principled response to Pereboom's four-case argument (Pereboom (2001) 110-120). As most readers familiar with the free will and desert literature know, Pereboom's original famous four-case account of manipulation involves: (i) moment-by-moment control through radio-technology; (ii) initial determination of psychology that causes all subsequent psychology; (iii) restrictive childhood indoctrination; (iv) ordinary socialization. Only case (i) bypasses proper mental functioning and undermines autonomy.
18. Perhaps the metaphor for defending Hell should have been not a strenuous uphill climb but keeping one's balance during an arduous and steep decline.
19. Perhaps God can blame those he causes in the foreseeable but not intended sense. The doctrine of double effect is thus relevant here to account for divinely determined evils. Or perhaps God can condemn the wrong-making features in his creations just as any wrongdoer can condemn their own actions. On the other hand, maybe He can't blame his creations for they did nothing wrong. If God was actualizing the best world then what his creatures did was for the best.
20. Boonin writes:

The harm requirement maintains that for a certain treatment to count as punishment, it must harm the recipient. But it is neutral on the further question of whether or not being subject to a harm might produce beneficial consequences in the future, including beneficial consequences that are great enough to outweigh (and perhaps even to justify) the immediate harmful ones. (Boonin (2008), 7)

21. A referee made this objection.

22. A referee wondered how a debt theory can also resolve the tension between justice and mercy. Many claim that justice should be tempered with mercy but this might suggest that they are at odds. They should be compatible rather than operating at the expense of each other. This is especially a requirement for theological ethics with its emphasis on God's mercy as manifestation of his loving nature and his loving and just nature being one and the same. Admittedly, there is traditionally considerable tension between justice and mercy. If justice is a virtue, and mercy means not bestowing justice, then it would seem that mercy is a vice. Such an unwelcome conclusion usually assumes as a premise a retributivist account of justice where the criminal deserves a certain level of punishment, and anything less is a miscarriage of justice. However, if punitive justice is determined by what restitution requires, then the forgiving victim is free to mercifully accept less compensation as a means to restoration and reconciliation than that typically demanded by the law. He can claim that the apologetic and remorseful criminal owes him nothing else, his debt either paid or forgiven, and can thus be restored to society as an equal. Justice as restitution and reconciliation allows both  $x$  amount of compensation and  $x$  minus  $n$  compensation. Justice doesn't demand either. It depends upon what the victim requires to be restored and reconciled to the 'release' of the wrongdoer. Perhaps this involves his moral worth being recognized if divine or human, though only the latter need their peace of mind regained and material wealth recovered. Therefore mercy in a restitutionist account is not internally at odds with justice. There is no need to invoke a value external to justice in order to reduce the debt the unjust owe.
23. Walls argues that Protestants can accept a 'sanctification' rather than 'satisfaction' account of Purgatory. C. S. Lewis is a well-known representative of this view. So Purgatory is not limited to just Catholic eschatology (Walls (2011), 59–91).
24. Peter asks: 'Lord, how many times shall I forgive my brother when he sins against me? Up to seven times?' Jesus answered: 'I tell you, not seven times, but seventy times seven.' (Mathew 18: 21–22). It is safe to assume that was not an order to offer 490 pardons but a command to forgive without limit.
25. This refusal to seek forgiveness and the parole analogy help resolve Sider's (2002) vagueness-based objection to Hell. Since the alternatives are just two, Hell and Heaven, Sider maintains vagueness brings a problem incompatible with God's justice: 'Any just criterion must judge created beings according to a standard that comes in degrees, or admits of borderline cases; but no criterion can remain simultaneously just – or at least be non-arbitrary – and consistent with the nature of a (dichotomous) afterlife' (Sider (2002), 59). So any cut-off line between eternity in Heaven and eternity in Hell would be arbitrary and unjust. To illustrate, consider Sider's claim of someone deserving Hell if he curses  $N + 1$  times. So let's say that  $N$  is 999 curses and so someone who curses 1,000 or more times deserves Hell. But could someone who uses profanity only one more time than the person in Heaven who uttered 999 swear words really deserve an eternity in Hell for the single additional expletive? However, if one user of profanity is repentant, and the other is not, the arbitrariness and unfairness and injustice of disparate punishments go away. One clearly wants to be forgiven, the other doesn't. The threshold will not be 1,000 curse words but the distinction between atoning and refusing to do so. Sider might reply that vagueness will infect the acts of contrition, reform, and reconciliation. So the arbitrariness of a cut-off will return as the above comes in all sorts of degrees. However, God would demand a *determinate* sufficient debt payment, an *unequivocal* apology and *definite* reconciliation. An eternity of indecisive contrition doesn't warrant Heaven. Even if it is not arbitrary to distinguish the repentant from the unrepentant, it might still be contended that there's still a vague borderline between minor sins that one never bothered to repent for that don't warrant Hell and major sins that at least, if unrepentant, do warrant Hell. Even minor sins would transform into major ones when sinners, confronted, refuse to ever apologize for them. That would be egregious pride and endless disrespect.
26. It is worth pointing out that libertarians still must confront a version of this problem, even if it is less disturbing. They need to explain why God didn't provide us with more inducements to virtue or fewer unruly desires for a libertarian will to confront and overcome. Molinists need to explain why God didn't just actualize worlds where no one would choose acts leading to damnation.
27. I hope that is not the case and so Hell is empty, though Purgatory will be quite full. But if an all-powerful God can't make a natural world without disasters, why should we be so confident that he can make a psychological world without imprudent obstinacy?