

## Evil, Freedom and Heaven

Simon Cushing

### 1. Introduction

My entry point into the topic of this chapter came while teaching the problem of evil in Philosophy of Religion, where I raise what I have called, for lack of a better name, the “Heaven Dilemma.” The problem of evil is, of course, why there should be any evil (and more broadly, suffering) in this world if it was created and overseen by a God who both has the power to prevent the evil befalling his children, and an all-loving nature that should prompt him to use that power. Those who believe the problem can be solved without abandoning the view of God that seems to give rise to it overwhelmingly appeal to free will.<sup>1</sup> The “free will defense”<sup>2</sup> relies on two core principles: that free will unavoidably bring with it evil, but that free will is also of such immense value (whether in itself or because it is necessary for some other great good) that it is better we have it than that there be no evil. But any version of the free will defense confronts the Heaven Dilemma as follows. If we assume both that existence in heaven is superior to our earthly existence in this vale of tears and furthermore that those in heaven commit no evil then the dilemma is this: either those in heaven are free without committing evil (which would undermine one of the core principles of the free will defense) or they are not free (which would undermine the other). Whichever it is, you can’t have both principles of the free will defense together, so the defense fails.

That’s not the end of the story, however. James Sennett<sup>3</sup> has suggested a way out of the dilemma by arguing, effectively, that freedom functions differently on Earth and in heaven such that the heavenly variety does not involve the risk of evil, yet is still of

comparable value to the earthly kind that does. An obvious pitfall of such an approach is that if it is possible *anywhere* to be free in a valuable way without the risk of evil, then the problem of evil recurs in a new form: why don't *we* have that *good* kind of freedom on Earth? Or, if it can be convincingly argued that you cannot have that kind of freedom on Earth, then the value of the risk-free heavenly freedom raises the Problem of the Apparent Pointlessness of Earth: why not have an entirely heavenly existence? Especially since we're going to spend an infinity there anyway (if we're lucky) so that the section of our existence spent on Earth is effectively zero. I examine Sennett's argument (and the related conception of freedom that it presupposes, argued for by Robert Kane) and possible responses available to him to rebut both these criticisms. I conclude, however, that the responses are unsatisfactory for a variety of reasons (infants who die before developing the capabilities of choice present a particular challenge) and that therefore theistic beliefs about heaven cannot be reconciled with the free will defense.

## 2. Freedom and the Problem of Evil

The following claims are central to theism, the view of God at the core of Judaism, Christianity and Islam:

Claim 1: OMNIPOTENCE: God is all-powerful.

Claim 2: OMNIBENEVOLENCE: God is all-loving.

Claim 3: CREATION: God created the universe.<sup>4</sup>

As I said in the introduction, these properties have been taken to be at odds with the quantity of suffering evident in the world. In response, the majority of theists endorse some version of the free will defense, and as such assert the following:

Claim 4: EARTHLY FREEDOM: Some of God's creatures, humans in particular, have free will.

This claim, which is rather vague (as there are many different ways to understand "free will"), is not enough for the theist by itself, because it needs to be shown both that it is entirely right and proper that God create such free beings, but that further, God could not do so without allowing them to produce the evils of the famous problem. Each of these assertions requires its own principle, and all variants of a free will defense assert some version of the following:

Claim 5: IEF (the Inevitability of Evil accompanying Freedom): Freedom is such that it is not in God's power to create genuinely free beings like those alluded to in EARTHLY FREEDOM without them performing some evil acts.<sup>5</sup>

Claim 6: TVF (the Transcendent Value of Freedom): The exercise of freedom in the right way is necessary to the creation of some phenomenon<sup>6</sup> whose value is such that it can outweigh evil that may result from the exercise of freedom in the wrong way.

Essentially, all versions of a free will defense<sup>7</sup> share the idea that given certain facts about our universe, even an omnipotent God was not able to give humans the kind of freedom that

is of transcendent value without evil resulting, but that God was still right to grant that freedom because any world that he could have created without freedom *or* evil would not be as good.

### 3. The Heaven Dilemma for the Free Will Defense

Heaven for theists is very like communism for Marxists: it's crucially important that it be coming, but very little is actually said about it. However, traditional theism is committed to the following three claims about the nature of heaven. First, some, if not all, humans can, under certain circumstances, enter heaven (whereupon they can be known as "the redeemed"). That this should be so is vital to heaven's roles in theism. Besides being a reassurance that one's existence does not end at the grave, heaven provides both an incentive to live one's earthly life right, and a salve against the injustices of this mortal coil. For heaven to function thus, heaven has to be both free of evil and better than Earth. Thus:

Claim 7: ACCESSIBILITY: Heaven contains<sup>8</sup> some "redeemed" ex-Earthlings.

Claim 8: IMPECCABILITY: The redeemed cannot do evil.<sup>9</sup>

Claim 9: SUBLIMITY: (Existence in) heaven is the best possible state of existence for its inhabitants.<sup>10</sup>

So much is, I think, uncontroversial. The trouble is, it appears that these claims taken as a group, alongside the evident truth that there is evil here on Earth, appear to have contradictory implications. Thus the Heaven Dilemma: one and only one of the following logically contradictory statements can be true:

Claim 10: HEAVENLY FREEDOM: At least some, and possibly all of the redeemed in heaven are capable of free actions (have free will).

Claim 10:\* HEAVENLY UNFREEDOM: None of the redeemed in heaven have free will.

As these two are contradictories, any set of claims that implies *both* of them must be inconsistent. But this is the case with the theistic claims we have already given. For it should be clear that IEF + IMPECCABILITY (claims 5 and 8) together imply HEAVENLY UNFREEDOM (10\*), because if freedom unavoidably brings evil (5) but the redeemed do not do evil (8) then they cannot be free (10\*). On the other hand, EARTHLY FREEDOM + TVF + SUBLIMITY (claims 4, 6 and 9) together imply HEAVENLY FREEDOM (10), because if we are free on Earth (4) and any state of existence with free beings in it is better than one without (6), yet heaven is better than Earth (9), then the fate of the once-Earthlings in heaven must include that valuable freedom they used to have (10).

If we remain committed to the claims about heaven, then the implication is that the two core claims of any free will defense cannot both be true at once. That is, either people in heaven are free, in which case it is false that freedom necessitates evil (IEF is false) or people in heaven are not free, in which case it is false that freedom is required for world to be optimal (TVF is false). What can the theist say in response?

#### 4. Unacceptable Response: Evil in Heaven

The simplest response is to discard one of the claims, in the hope that the remaining claims are mutually consistent. Of course, I contend that a true theist is committed to all of

them, but of all of them perhaps IMPECCABILITY is least obviously essential. If we discard it, however, what are we saying? There are two clear alternatives.

The more extreme alternative is to assert that, not only is evil possible in heaven, it will happen or has already occurred. Some support for this position might be found in the story of Lucifer rebelling against God. However, apart from the fact that this story is not necessarily canonical,<sup>11</sup> Lucifer's presence in heaven was not achieved in the same way as one of the redeemed and we are focusing on the state of being of (ex-) humans in heaven. Furthermore, the existence of evil-doing by the redeemed in heaven would open a big enough theological can of worms to merit ruling out this response for the purposes of this paper.<sup>12</sup>

The milder alternative, then, is to allow that evil in heaven is in fact *possible*, even if it is not, and never will be, *actual*. That is, instead of IMPECCABILITY we have:

Claim 8: ABSENCE: There will not be evil in heaven (even if it is always possible).

There are reasons both philosophical and theological why this initially plausible response will not do. I'll let Sennett handle the theology:

[T]raditional theism regards heaven as a place in which evil is completely eradicated—it is not even *possible* that any should arise. The difference between heaven and Earth is not simply that Earth contains evil while heaven does not... If heaven is nothing more than a place where the possibility of freedom and no evil is realized, then the absence of evil in heaven is purely contingent on the choices of human beings, and not a matter of God's sovereignty or the nature of heaven at all. But the traditional view of heaven in that it owes its purity to the unmediated presence

of God. Heaven is *essentially* pristine, grounded in divine immanence, not contingently so due to the fortunate choices of humans. (69)

Furthermore, if heaven is only evil-free contingent on the choices of its human occupants, then it is constantly in danger of losing its evil-free status, since it is always in the power of those occupants to introduce evil into heaven. But certainly the idea that heaven might yet become a place of sin and rebellion is contrary to traditional theism. In order for heaven to be essentially pristine and free from future corruption, it must be necessarily evil-free—it cannot be possible for there to be evil. (70)

The philosophical problem with replacing IMPECCABILITY with ABSENCE is that it appears directly to undercut IEF, because it allows that a whole realm of being can contain agents acting freely (for an infinite time, no less) without evil resulting. This raises a new version of the problem of evil: if it is possible for freedom to exist without evil *in heaven*, why are the two inextricably linked on Earth? Allowing freedom to exist without evil *at all* seems to lay the responsibility for the existence of evil at God's feet: he could have made us genuinely free without the accompanying cost of evil.

Another way to think of this problem is as an instance of a more general problem for views that incorporate the notion of heaven: the problem of the Apparent Pointlessness of Earth (henceforth APE). If we can enjoy in heaven every possible good that Earth brings, along with incalculably many more, and furthermore avoid the pitfalls of sin (which are usually taken to include the possibility of an eternity in hell, should the sin be serious enough), then why should any time spent on Earth be of value? Aren't the infants who die before they have a chance to sin obviously better off in all respects than those who don't,

especially those who go on to commit heinous sins on Earth and condemn themselves to hell?<sup>13</sup> To make APE vivid, one should try to take the point of view of God facing the choice of bringing into being just heaven, or heaven plus an extra, very inferior, evil-ridden plane of existence. One might by analogy imagine choosing between building the Taj Mahal, or a slightly smaller Taj Mahal with a tin shed tacked on the side. This will come up again below, but for now I conclude that discarding IMPECCABILITY is an unacceptable response to the dilemma.

Sennett believes that there is a way out of the dilemma that draws on a sophisticated conception of the nature of free will that has been developed by Robert Kane, a contemporary specialist on the topic. So to make sense of Sennett's attempted solution we need to take a detour into some background about the debate over the nature of free will.

## 5. Compatibilism vs. Libertarianism and Kane's Compromise

There are two major camps among philosophers who believe humans have free will. The dominant position is *compatibilism*, the view that even if one's actions are fully determined by the state of the universe and the laws of nature, one's actions can still be free.<sup>14</sup> According to classic compatibilism of the kind defended by philosophers like Thomas Hobbes, David Hume and John Stuart Mill, an action is performed freely simply if the person performing it wanted to do so and was not coerced against her will.

Compatibilism is a non-starter for free will defenders, however.<sup>15</sup> As Kane notes, it is ill-equipped to provide the firewall between God and responsibility for evil that they desire:

Compatibilists believe that freedom (in all the sense worth wanting) could exist in a determined world. But if we did live in a determined world and it was also true that



God had created that world, then everything that happened in that world would have been predetermined, and hence predestined, by God's act of creation. The ultimate responsibility for all that occurs would go back to God.<sup>16</sup>

Free will defenders thus opt for the other major position on free will: metaphysical *libertarianism*, which rejects both compatibilism and determinism. As we are interested in the variant of libertarianism that Sennett defends, let us begin, as he does, by citing a favorite argument of contemporary libertarians against any compatibilist position, what Peter van Inwagen calls the "consequence argument:"

If determinism is true, then our actions are the consequences of laws of nature and events in the remote past. But it is not up to us what went on before we were born, neither is it up to us what the laws of nature are. Therefore, the consequences of these things are not up to us.<sup>17</sup>

This argument sets a necessary condition for an action to be "up to" the person performing it: there must be a break in its causal history such that what follows the break is not determined by what came before it. Libertarians believe they can give an account of freedom that meets this condition. Exactly where the break needs to occur, however, is a matter that divides libertarians.

Let us crudely divide the universe into three elements, relative to any particular individual. The first is everything (both in space and in time) that is "external" to that individual, such as environmental factors and everything beyond the control of that individual. The second is all that might be called "internal" to the individual, which comprises what you might call her "self" or her "character:"<sup>18</sup> things like beliefs, desires,

predilections, temperament, virtues, vices, *et al.* The third is the events that (for lack of a better term) result from the individual, which can rightfully be attributed to her. Most relevantly for our purposes, these would include all her actions. Libertarians could demand a break between the first and the second element (call that 1B2 Libertarianism, where the “B” represents the indeterministic break) or the second and the third (2B3 Libertarianism).

What you might call “classic” libertarianism is of the 2B3 variety. On this view, for every free action, given a complete description of the makeup of an individual’s character and the circumstances, the resulting action is still undetermined. A popular way of conveying this idea is to imagine two distinct possible worlds<sup>19</sup> that are identical to the last detail right up to and including the character of the person acting, but diverging at the moment of choice, so that in one world the agent does one thing and in the other she does another, but in each of which the action was performed freely.

Putting the break between character and action in this way has long been controversial. Mill pointed out that we would be insulted if our friends were not able to predict what we would do in most situations.<sup>20</sup> And Hume argued that, without a deterministic link between character and action, an agent’s actions cannot be a source of blame or credit to her.<sup>21</sup> As we have seen, compatibilists insist that what we look for in freedom is freedom from coercion. More broadly, they insist that to be free is to be in control of one’s actions. But the control we look for *requires* that our character determine our actions. A break of the kind this classic libertarianism requires would *impede* freedom, not facilitate it: the picture it presents looks more like random things happening to people rather than people in control of their actions. Call this compatibilist charge against libertarianism the “Indeterminism Undermines Control” (IUC) complaint.

Kane allows that the compatibilist may be on to something, in two respects. First, indeterminism between character and action is a barrier to freedom and not the source of it:

One thing does seem to be true about control which critics of indeterminist freedom have always maintained: indeterminism, wherever it occurs, does seem to diminish rather than enhance agents' voluntary control... Moreover, this limitation is connected to another, which I think we must also grant—that indeterminism, wherever it occurs, functions as a hindrance or obstacle to our purposes that must be overcome by effort.<sup>22</sup>

Second, it seems right to say that we can have natures that both determine, and rule out alternatives to, certain actions that are nonetheless free. Both Kane and Sennett endorse the example introduced by (compatibilist) Daniel Dennett of Martin Luther. Luther supposedly said of his famous action, “here I stand; I can do no other.”<sup>23</sup> Dennett contends that we should take this seriously as an example where Luther's character was such that *no other action was possible*.<sup>24</sup> In other words, his character determined his action. Dennett further argues that Luther's act was nonetheless free. Kane agrees on both counts, which means that he cannot be a classic libertarian of the 2B3 variety. However, Kane keeps his libertarian bonafides by instead endorsing a kind of 1B2 view: actions that are determined by one's character can nonetheless be free, but only so long as that character has not been entirely determined.

So to recap: compatibilism is a view of freedom that makes free actions entirely predictable to an omnipotent being and which would allow that being to create free beings who will never commit evil. It thereby undermines IEF. However, it faces the consequence argument, which demands that there be a point in the history leading up to an action such that what follows the break is not entirely predictable even on the basis of total knowledge of the universe up to that break. Libertarianism claims to meet the challenge of the consequence argument, and also open the door to IEF.<sup>25</sup> However, libertarianism faces IUC, the challenge

that the break it demands simply undermines the control that is a core feature of our conception of what it is to act freely. IUC, if telling, would undermine *both* claims of a free will defense. TVF would fall, because, if no praise or blame can accrue to the character of the agent, what is the value of the act? And with no blame, where is the evil that IEF asserts?<sup>26</sup>

Kane argues that for a freedom that is truly valuable we must somehow avoid both the consequence argument and IUC, and believes that his version of 1B3 libertarianism can do that. There is one more element to add, however, because it should be clear that simply having an indeterministic break between external forces and one's internal character is, by itself, not the magic bullet that renders one free. To see that, consider what I'll call the "Marvel Origin Case." Suppose that we are considering two individuals: one, Rick, has a character that (a) is entirely determined by external forces, and (b) *à la* the Martin Luther example, entirely determines his actions. Another, Bruce, also has a character that determines his actions, but unlike Rick, a key aspect of Bruce's character was formed as a result of some bizarre neurochemical accident involving the indeterministic decay of some radioactive substance he was exposed to *in utero*, thereby providing the indeterministic break that ensures that his character trait is not a "consequences of laws of nature and events in the remote past." In such a case, the requirement of the consequence argument is met, yet this surely would not be sufficient to make Bruce's actions valuable in the IVF sense if Rick's are not. We need more. Kane realizes this, and suggests that the extra element is that the character of a person whose determined actions can nonetheless be free is *self-formed*. He argues that there are *self-forming* actions<sup>27</sup> (SFAs) that are distinct from regular actions because their effect is to shape one's own character, and it is only if one's character has been (at least partially) shaped by *undetermined* (i.e., libertarian rather than compatibilist) SFAs that the trick is turned and you get the ideal form of freedom that avoids both the consequence argument and IUC:

All actions done of our own free wills do not have to be undetermined SFAs of this kind... But if no actions in our lifetimes were of this undetermined self-forming or will-setting kind, then our wills would not be our own free wills and we would not be ultimately responsible for anything we did.<sup>28</sup>

There is more to be said, and we'll get to it, but that outline is enough to prepare us for Sennett's use of a Kanean conception to attempt a solution to the Heaven Dilemma.

#### 6. Sennett's Solution: Libertarianism on Earth, Compatibilism in Heaven

Sennett broadly endorses Kane's view, as I have said. As he puts it, "A character that is libertarian freely chosen is the only kind of character that can determine compatibilist free actions."<sup>29</sup> He also adds some helpful terminology: actions like Martin Luther's are "proximately determined" and can be free, so long as they are "remotely *undetermined*." With these terms in hand, the essence of Sennett's response to the Heaven Dilemma is given by the following passage.

Consider the conception of general freedom. A world segment includes general freedom just in case that segment includes libertarian freedom or (proximate) compatibilist freedom. It can be argued, consistent with the free will defense, that it is general freedom, not libertarian freedom *per se*, that manifests the freedom good. Either libertarian freedom or (proximate) compatibilist freedom is sufficient for general freedom. Thus, any world segment excluding libertarian freedom but

including (proximate) compatibilist freedom is a world segment manifesting the freedom good.<sup>30</sup>

Let us spell this out in steps:

- i. IEF applies to any actions that are free in a libertarian sense, but not to any actions that are proximately determined.
- ii. Our actions on Earth must include actions that shape our characters, and such actions are only *truly free* (that is, valuable in the TVF sense) if they are free in a libertarian sense (and if so they are Kanean SFAs).
- iii. From i and ii it follows that on Earth, in order to get the “freedom good,” you must have actions that also bring evil.
- iv. Our actions in heaven, however, *do not* include actions that shape our characters, because our characters are *fully formed* before we enter heaven. Instead, all heavenly actions are “actions *from* character” (AFCs).
- v. AFCs *do not* have to be free in a libertarian sense to be valuable in the TVF sense: a proximately determined AFC can still “truly free” and thus instantiate the freedom good so long as it is also remotely undetermined.
- vi. Evil is entirely avoidable when performing proximately determined AFCs (even though such actions are part of “general freedom”), so long as the character that proximately determines the action is morally pure.
- vii. From iv-vi it follows that in heaven we can get the freedom good without evil.

Sennett’s attempted solution to the inconsistency of our ten theistic claims is, in effect, a modification of claims 5 and 10. Claim 5, IEF, becomes

Claim 5.1: IELF (the Inevitability of Evil accompanying **Libertarian** Freedom):

Libertarian Freedom is such that it is not in God's power to create genuinely free beings like those alluded to in EARTHLY FREEDOM without them performing some evil acts.

Claim 10: HEAVENLY FREEDOM, becomes

Claim 10.1: HEAVENLY COMPATIBILIST FREEDOM: All acts in heaven are proximately determined. Those members of the redeemed (possibly all) who act freely do so because their acts are also remotely undetermined.

Does this strategy allow Sennett to avoid APE? Consider the following thought experiment.

## 7. Eartha and Celeste

Two beings with identical characters arrive simultaneously in heaven. One of them, Eartha, had a pre-existing life on Earth during which time she underwent SFAs, while the other, Celeste, was simply brought into existence in heaven, character fully formed. Celeste's nature is in no way derivative of Eartha's: it is mere coincidence that they have identical natures.<sup>31</sup> Imagine, for now, that heaven is such that one can observe its inhabitants performing actions. We observe Eartha and Celeste both performing the same action (adoration of God, say), for the same reasons. In both cases the actions are proximately determined AFCs. Are both equally valuable acts? On Sennett's view they cannot be, for if they were, then Eartha's earthly existence made no difference, and APE has not been rebutted. And in fact, according to Sennett, although both Eartha's and Celeste's are identical actions of beings that are identical in all respects that can be measured at the time they act,

only Eartha's act of adoration is the kind of truly freely-given love that could make it valuable.

Thus understood, Sennett's view provides the following answer to APE. Earth has essential value because the only way for determined choices in heaven to be free in any valuable sense is if the character determining them was "self-formed" by the agent in her previous life on Earth. Thus Earth's existence, evil-ridden as it is, is essential to ensure freedom in heaven.

Does this seem a defensible analysis of Eartha and Celeste, however? According to Sennett, we can look at Eartha and remark, "ah, love given freely—is there anything of more value in the universe?" then turn to observe Celeste's *identical* choice, performed for *identical reasons*, as a result of an *identical* character, shake our heads and say sadly, "what a worthless, puppet-like charade."

It gets worse: suppose we have no knowledge of their history prior to being inhabitants of heaven (and they're no help, because they both have the same "memories") but are told that only one of them has an earthly history, but not which one. We could scan them down to the subatomic level (or whatever celestial substance of which the redeemed are composed), find them identical, and still have to say "well, one of them's a puppet, while the other is a free agent exhibiting the freedom good that has value beyond price, but there is absolutely no way for us to tell which."

Perhaps a defender of Sennett's account might make the following response: "Such a distinction as that being made between Eartha and Celeste shouldn't be so surprising: we can imagine a human and an android performing identical tasks, but we would only be inclined to say that the former is free." But that would be a faulty analogy. For one thing, the reason why we don't think of an android as free is because we assume it is not conscious, but Celeste is as conscious as Eartha. More importantly, in the human vs. android case, the distinction



would presumably be made on the basis of the human having and the android lacking *the capability* to choose otherwise. But this is something that Eartha and Celeste lack equally.

Sennett might instead respond that the idea that whether or not an act is free in the way that is valuable depends in part on facts about history should not be as counterintuitive as I seem to find it. For comparison, questions of personal identity are usually taken to be similarly history-dependent.<sup>32</sup> Who is my sister and who is her Star-Trek-transporter-malfunction-produced duplicate cannot be settled simply by examining their current makeup, because it is identical. It depends on which of them originally entered the booth and pressed the button. Or, to press the analogy with “self-forming” actions, my real sister’s memories *really are* her memories because *she* was the one who did the action being remembered (and “formed” the memories in the first place), whereas the “memories” of the duplicate are not hers—she is not entitled to them.

But actually, this case gives grounds for a pushback against the claim that it makes sense to call Eartha free while denying that Celeste is. Imagine two possible worlds: in one, my sister survives but the clone dies. In another, the reverse happens. In both worlds, the surviving individual performs exactly the same acts, none of which, let us stipulate, are undetermined in the way that SFAs have to be. My intuition is that if, in the world where she survives, my sister acts freely, then in the world where the clone survives *she acts just as freely*. But this would not be so on Sennett’s view, because the clone’s character was not formed by the clone, but by my sister (just as Celeste’s character was formed by God, and not by Celeste).

I realize that I’m just appealing to intuitions here, and I am aware that others might have no problem with the idea that one’s freedom now cannot be assessed without knowledge of the past. So it behooves me to look more closely at the key notion of SFAs on which Sennett’s strategy rests.

## 8. Self-formation

To that end, the following questions need addressing. First, what exactly is the “self” or “character” (or occasionally, “will”) that needs to be formed on Earth and perfect in heaven? And second, what exactly is it for an action to be “self-forming” and how is that supposed to make the difference between possessing and lacking the capacity for determined yet free acts?

Let us use “self” to stand for all and start from the rough idea that it is what we can point to as explanation for what we contribute to our actions. However, that crude outline encompasses a spectrum of concepts from “thick” to “thin.” A thick concept of a self would include just about every belief, desire and preference that any individual possesses.<sup>33</sup> The “thicker” a self is, the more likely it is to include elements that cannot plausibly be said to have resulted from choices of mine.<sup>34</sup> Certainly deeper tendencies, like (perhaps) an addictive personality, can be ineluctable and enduring elements of a thick self.<sup>35</sup> This appears to indicate that a thick self will contain elements that are not “self-formed.” I believe this to be so, and we will explore the implications below (as well as a thinner conception of the self).

Now let us look at *self-formation*. According to Kane and Sennett, having some SFAs in one’s past marks the difference between free proximately determined acts and those that are mere facsimiles of freedom. What is it about self-formation that makes this difference? I think the clearest answer follows from a well-known libertarian criticism of classic compatibilism, as follows. Recall that classic compatibilism suggests that freedom is about getting what you want without outside interference. That is, it is *other people* who stop you being free, but if you’re getting what you want, you are unimpeded and thus, free. Libertarians pointed out, however, that the two parts of classic compatibilism (getting what

you want, and not being controlled by others) are not coextensive. What if those other people, rather than force you to do the things that they want you to do against your will, instead manipulate your character so that you now *want what they want*.<sup>36</sup> So, perhaps I brainwash you to want to give me all your money. Even if you give it to me “of your own free will,” and insist you are happy to do so, the vast majority of us would agree that this is not true freedom, certainly not the kind that underpins responsibility. Thus (the libertarians continue) it would seem that there is a strong case that, for an agent’s actions to be free, her character must not have been manipulated by others. From this negative requirement Kane gives us his positive solution of SFAs: if instead that agent’s character was formed by *herself*, then this would be the deciding factor that made her actions truly her own, ensuring that she be acting of her own free will. Sennett appears to concur:

What I have in mind is this: many of the character traits we display—honesty, for example...were formed in us as a result of consistent behavior patterns that developed into habit. These behaviors were not always habitual, but began as overt, deliberate actions, perhaps taken after not a little pondering and soul-searching... Perhaps she now is so practiced in the art of probity that she responds with ingenuousness and veracity without hesitation or forethought. Her character demands and determines that she do so. But she could not have reached such a state had she not deliberately chosen honesty from among genuine alternatives in the past.<sup>37</sup>

There are two problems with this explanation of the value of SFAs to freedom. The first is that more sophisticated compatibilists *agree* that brainwashed acts are not free, and that self-formation is important, and have *fully compatibilist* accounts of self-formation.<sup>38</sup> The second is that, given Kane’s expressed concern over IUC, adding indeterminism to these acts of self-

formation (as Kane and Sennett both want in order to avoid the consequence argument,<sup>39</sup> and Sennett needs to support IEF on Earth) appear to *undermine* them, not complete them.

Sennett is aware of the possibility of compatibilist SFAs, as he makes clear here:

Of course, this scenario is open to the charge that these precedent deliberate actions were themselves the determinate product of her character as it stood at that time, and so the standard compatibilist position would insist. But again such a response would suggest a regression that will inevitably lead back to before her birth, and the Consequence Argument again threatens.<sup>40</sup>

Again, though, his response is simply that the consequence argument requires indeterminism. But the Marvel Origin Case shows that simply adding indeterminism is not enough to make the difference between freedom and its lack, so you can't just stick a bit of indeterminism onto something that doesn't need it and think it fixes the problem. It has to be that indeterminism is *baked in* to SFAs (in a way that it is *not* in regular actions). Sennett is no help here, but Kane, who appears to take IUC much more seriously, has attempted an account that does this.<sup>41</sup>

## 9. Plural Voluntary Control

Let us bring out the real contrast between compatibilists and libertarians over the role of indeterminacy with an example Kane himself uses.<sup>42</sup> Imagine two possible worlds, identical up to a moment of choice. In one, a businesswoman chooses to stop and help a person in need, in the other she hurries past to get to a meeting. The libertarian must insist that both different futures must be entirely physically possible consistent with the same past for it to be

the case that the businesswoman's choice (of either future) is truly free. Compatibilists are simply incredulous at this, charging that if her character (plus the laws of nature) doesn't favor one over the other then *neither* choice is truly "hers" and she has experienced a random event rather than made a decision that is a result of deliberation. Kane attempts a compromise.

His first move is to argue that indeterminism does not *completely* undermine responsibility if it is successfully overcome:

Consider an assassin who is trying to kill the prime minister, but might miss because of some undetermined events in his nervous system which might lead to a jerking or wavering of this arm. If he does hit his target, can he be held responsible? The answer... is "yes" because he intentionally and voluntarily succeeded in doing what he was trying to do—kill the prime minister. Yet his killing the prime minister was undetermined. (227)

This is something that the compatibilist can concede, because the freedom is entirely explained by the extent of the determinism, and diminishes the greater the role of indeterminacy. However, the gulf between the libertarian and compatibilist is still wide: classic libertarianism must allow that even if there is no indecision and the businesswoman character is entirely set towards helping, then she could still walk on by, or even do something horrendous, like pushing the victim under an oncoming car.<sup>43</sup>

Kane's second move is to reassure the compatibilist that these worse options will not happen. Instead, SFAs of the kind he envisages involve choices where the competing urges at work within the agent's character pull with essentially equal force. She is genuinely

undecided between them, but would endorse either of them as both result from an aspect of her character:

[W]e know the brain is a parallel processor and that capacity, I believe, is essential for the exercise of free will. In cases of self-formation, agents are simultaneously trying to resolve plural and competing cognitive tasks. They are, as we say, of two minds. But they are not therefore two separate persons. They are not disassociated from either task. (231-2)

According to Kane, it is the competition itself that produces the indeterminacy that satisfies the requirement of the consequence argument.<sup>44</sup> This, then, is Kane's notion of "plural voluntary control." "plural" because it involves a person being "of two minds," but "voluntary control" because, if we think of the businesswoman's "control center" overseeing her competing desires, it is saying "my second-order desire is that one of these competing desires be realized, but I will leave it up to chance to see which one." Her "control center" desires something (that one of the desires be realized) and that overall desire is itself realized. This seems like the definition of classic compatibilist freedom, but with indeterminism baked in to satisfy the consequence argument. Everybody wins, right?

Not so fast. This seems okay until we see the weight being placed on this notion of freedom. Acts like this, where there is plural voluntary control, are said to be *essential* to enable regularly determined acts to be free in the future. But to the compatibilists these actions are simply *inferior* versions of the determined acts. They are far from indicators of one's true nature, or cases where one made a real decision that changed one's life in a profound way. These are like mental "coin-tossings," cases where the person *didn't care either way*, and *for that reason* were prepared to allow randomness to choose. This is an

*abdicating* of responsibility, not an instance of *true* responsibility. Certainly not the kind of life-framing decision one expects in a SFA. Imagine somebody whose life was filled with nothing but such SFAs: I think we would call such a person aimless rather than admire her for the amount of genuine free self-formation she indulged in.

Even worse, while I have been talking of a “control center” that oversees the competing desires, this was just a metaphor to illustrate Kane’s view. Kane puts it this way:

[W]e say in effect, “Let’s try this. It is not required by my past, but is consistent with my past and is one branching pathway my life could now meaningfully take. I am willing to take responsibility for it one way or the other.” (238)

But this description is disingenuous. It smacks of the “homunculus” view of the mind, as a mini-me inside my head watching monitoring screens and making decisions. If the brain really is a “parallel processor” then there may be no such central control. Certainly in the businesswoman case she is unaware that she is allowing a random process to control her fate; it just happens. There is no second-order desire. Ironically, if she were actually tossing a coin, she would be more responsible, but we would be very hesitant to call that worthy of a self-forming action. I therefore conclude that Kane’s conception of plural voluntary control fails to offer an account that does justice to the notion of self-forming action

But even supposing I am wrong about this, and Kane’s account were a resounding success, it would not be of use to Sennett. Recall that Sennett needs his indeterminism to account for all evil on Earth. It must be that, despite the perfection of God’s creation, including every aspect of human natures that can be attributed to God, the nature of earthly freedom is such that all of the horrors readily apparent on Earth will result. But we saw that Kane was insistent that the indeterminism only resulted from a clash of (two or more) already

existing desires. This means that any evil that resulted was already there in the individual. If this is true even of that individual's *very first* SFA, then her self must already be corrupted *before she ever has a chance to form it*, which would lay the blame for any evil choices on her creator.<sup>45</sup> Or to put it another way, Kane's account means that IEF need not even be true of libertarian freedom. God could create us with such natures that even our undetermined choices could never result in evil. (Ironically, in an article broadly sympathetic to Sennett's account but wishing to expand the freedom of the redeemed, Timothy Pawl and Kevin Timpe argue for the possibility of a kind of freedom that is similarly both libertarian and free from the risk of evil, where the redeemed get genuine undetermined choice, albeit between only good options, like "the choice either to sing in the heavenly choir or to play the harp."<sup>46</sup> I have different issues with Pawl and Timpe's account that space does not permit airing, but suffice to say, if we can have a genuinely valuable, genuinely undetermined kind of freedom with no risk of evil resulting (provided God sets things up right), then we are once again confronted with the worry that there would not be evil if God were as theism says.) So it looks like Sennett needs somehow to extract Kane's notion of SFAs without his particular account of how they work. But Sennett offers no alternative, and thus his account stands or falls with Kane's.

## 10. Clay and Rock, Mary and Jezebel

Further problems arise concerning the as-yet-unsettled precise nature of the self that is to be formed. We have been working so far with a thick self, which, we have seen, includes many elements that cannot plausibly be said to have originated from free choice, and thus were "given to us," and may even stubbornly resist attempts by us to mold them (assuming we are even aware of them). In fact, it seems to me that some people are much more open to



having their characters shaped than others, so to what degree (if at all) “self-forming” is possible may vary across individuals in a way that they have no control over.

Consider potential self-formers Clay, who has a malleable character, and Rock, who has a supremely inflexible character. Suppose further that both Clay and Rock perform the same undetermined virtuous action (laying aside compatibilist qualms for now). With Clay, this action lays the groundwork for a reformed character. Truly, he is on the path to a perfected self. With Rock, however, this is just a one off. Tomorrow his undetermined action will be vicious, and he remains incorrigible. What should Sennett say about this difference? It seems like Clay might be on the way to freedom in heaven, while Rock is not, but only because of a feature of his self that was determined by forces beyond his control (presumably, for the theist, by God). Would Sennett say that Rock’s act, free though it was, and identical to Clay’s as it was, is *not* self-forming, precisely because it did not change him? That seems unfair, especially if, in the extreme, he never performs *any* SFA and is thus robbed of the chance of freedom in the afterlife simply because of an inflexible nature.

Alternatively, imagine Mary and Jezebel. Mary’s inborn nature is saint-like—bad thoughts never even occur to her—while Jezebel was born with a character riddled with desires, tendencies and even perhaps mental and physical disorders that render it very far from perfection. Nonetheless, somehow (let us stipulate) both perform the same set of undetermined SFAs (something that could only happen if the actions truly were undetermined by the natures of those performing them). Mary’s character is perfected in short order, and is thus ready for heaven, but Jezebel’s is not. This seems unfair, and in fact seems an injustice that cannot, by its very nature, be explained away by a free will defense, because Jezebel is disadvantaged over somebody who makes *exactly the same choices*.

It seems to me that there are two lines of response open to Sennett over cases like these. The first would be to argue that it is *effort* that counts and that Rock's effort will be recognized as much as Clay's, and Jezebel will be welcomed into heaven if Mary is.

The problem remains, however, that Rock and Jezebel have unperfected characters. Does this mean we are rejecting IMPECCABILITY? If so, we run the risk of evil in heaven. To block this, perhaps we can stipulate that, once you get into heaven your selves are perfected automatically by God, so that Jezebel acquires as pure a character as Mary. No; surely we *cannot* stipulate this. Apart from the fact that there would be issues of personal identity (in what sense is heavenly, perfect Jezebel *the same person* as flawed, earthly Jezebel?), this new Jezebel seems disturbingly similar to Celeste, in that the perfect self that will determine all her actions henceforth was *not* a product of her SFAs.<sup>47</sup>

Alternatively, perhaps Rock and Jezebel can keep their characters in heaven but be blocked from their choices ever producing evil. But this possibility cannot seriously be entertained by somebody who argues, as all free will defenders do, that you cannot thwart even the free choice of Satan to torture humans without robbing freedom of its true value.

The second line of response might be for Sennett to argue for anorexia of the self. That is, instead of a "thick" self that includes attributes beyond the control of its possessor, we should narrow down the self to consist simply of the things that *have* been chosen by that person. Perhaps we could reserve the name "self" for the thick self and "character" for this thin self.<sup>48</sup>

This maneuver again threatens personal identity. Are we to imagine that our "determined" earthly characteristics, like a love of beer, sex and coffee, will be replaced in heaven by more "appropriate" desires, albeit with no loss of self? If the response is that all that really matters about us, our essence, is our freely chosen character, then what about those who die very young (who have an even more serious problem, as we shall see in the next

section) who would seem to have *no characters at all*, not having had the chance to perform any SFAs? What would they be like in heaven, with such “thin” characters? (Remember, Sennett cannot allow them to expand their characters in heaven, because SFAs are confined to Earth.)

Second, even given this maneuver, surely it is clear that those whose pre-determined natures and circumstances were such that their choices were simply between two “good” options would further be advantaged. It would be like comparing the portfolios of two artists, one who had unlimited resources and vast natural talent, and one who only had a twig to work with and no limbs to hold it.

Third, if we define character as what results from SFAs, then we are forced to say that Kane’s businesswoman 1 was good because indeterminism resulted in her helping, while businesswoman 2 was bad, *despite having exactly the same beliefs, desires and intentions*. This makes businesswoman 2 look merely unlucky rather than deserving of heavenly rejection.

Finally, remember that the “self” in self-forming is supposed to be the active former as well as the passive material being formed. But in this extremely narrow view of self-as-character, how do we explain the first SFA? Before an act has occurred, there is no character to do the forming *or* to be formed. Who exactly is it *performing* the SFA? On this view, we would all bring ourselves into being out of nothingness. And that is beyond even God.

## 11. “Little Angel Automatons”

Let us imagine that there is an account of SFAs that can solve all of the foregoing problems, so that a first SFA is possible, and that even the first SFA of an unblemished character could result in evil. Nonetheless, the fact that there has to *be* a first SFA creates a

huge problem for Sennett. To introduce this problem, it is also worth reminding ourselves of what started this whole discussion: the problem of evil. A phenomenon that has made the problem particularly vivid for me since I first read about it while teaching Bioethics is the horrific genetic condition *Hallopeau-Siemens syndrome*. It was after Sanne, a baby with this condition was born on his ward, that Dutch pediatrician Eduard Verhagen began to advocate for euthanasia for infants:

Her skin would literally come off if anyone touched her, leaving painful scar tissue in its place. The top layers of mucous membranes inside her mouth and esophagus fell away any time she was fed, which was done by tube... Her cry was not that of a normal, healthy baby but the shriek of an extraordinarily sick one. And her vital signs—heartbeat, blood pressure and respiration—reflected those of a child in extreme stress. Pain relievers seemed to be useless. Making matters worse, Dr. Verhagen and his colleagues had to bandage Sanne's scar tissue knowing they were contributing to a vicious circle: every time they replaced the bandages, a little more skin fell off... Her parents demanded an end to her suffering, which moved Dr. Verhagen to consider euthanasia. Fearing criminal prosecution, he and hospital officials refused and eventually sent Sanne home, where she died of pneumonia half a year later. Dr. Verhagen felt he had failed Sanne and her parents, believing all three had suffered longer than necessary.<sup>49</sup>

Surely the parents of any baby born with this condition are caused to wonder why God would allow it. This is what we need theodicies for. Of course, it also reveals a weakness for the free will defense: it is particularly ill-equipped to deal with so-called “natural” evils like this, because they look so clearly like “acts of God.”<sup>50</sup> Nonetheless, the free will defense is by far

the most respected response to the problem of evil, so it is the best answer we will get. For the purposes of this paper, though, cases like Sanne's present a different challenge. We are asked to believe one of two unacceptable things: either that such an infant is able to perform SFAs some time in her short agony-filled (or heavily sedated, if she is lucky) earthly existence, or that she does not, and is thereby condemned not just to earthly agony but to an afterlife of unfreedom.

The first option is untenable. Kane allows that SFAs can be rare.<sup>51</sup> On any plausible account of self-formation, SFAs presumably require a certain level of brain maturity. We do not regard young children as responsible moral agents for good reason, and we *certainly* do not regard infants as such, so why should they be capable of true self-formation? There are quite rightly no human rights organizations arguing for the freedom of infants to make their own life decisions, and nobody regards paternalism towards them as an intolerable restriction of precious-beyond-price freedom.

So it must be that babies like Sanne die before performing even a single act of self-formation. On Sennett's account this means they end up just like Celeste. Her every action in heaven is unfree, however indistinguishable it would be from those of supposedly free members of the redeemed like Eartha. Call this the Challenge of the Little Angel Automaton. It is a serious problem indeed for Sennett's two-stage conception of freedom, because it means that *millions and millions* of the redeemed are denied precious freedom, through no possible fault of their own.<sup>52</sup> For their entire infinite existence they will never sample the good that is supposed to be beyond price, and worth the cost of all the misery on Earth.

What is worse, for Sennett this problem combines with APE to form an apparently inescapable dilemma: *either* you don't need Earth to get the freedom good, in which case we face APE and the attendant problem of evil (as well as an incentive to usher our infants into

heaven before they risk suffering or have the chance to sin and risk an eternity in hell), *or* every soul must be formed on Earth, in which case those millions who perish before having a chance are denied the chance at freedom itself, and all the things that require it, including, presumably, love of God in heaven. Can this new dilemma be avoided? In what follows I will assess various attempts to do so.

## 12. Possible Responses

### i) No Possibility of Death before First SFA

If there were no possibility of a person dying without performing a SFA then we can avoid a heaven full of Little Angel Automatons. But how would this work? It would require SFAs to happen essentially immediately a human came into being. But supposedly SFAs are instances of individuals having some control over their own personality, and hence their value:

if there were no such undetermined SFAs in our lifetimes, there would have been nothing we could have ever voluntarily done to make ourselves different than we are—a condition that I think is inconsistent with our having the kind of responsibility for being what we are which genuine free will requires.<sup>53</sup>

Of course, I have argued that Kane's own plural voluntary control account pretty much makes nonsense of the idea that SFAs are us "voluntarily" forming ourselves, but for the newly-formed zygote or even the late-term fetus, such a description is positively laughable. If we are to say that SFAs happen immediately a being comes into existence then we need a

different, weaker account of SFAs, and then we would also need an explanation of why they are of value. And for any weaker conception of a SFA, the challenge is to show that it is not simply an undetermined event that happens to an unknowing, unchoosing, pre-moral-agent, something the Marvel Origin Case showed was of no help to the libertarian. I conclude that this response is a non-starter. We should always remember that the “self” in “self-forming” is not just the thing being formed, but the agent doing the forming. Thus self-forming cannot happen unless there is an existing agent able to make choices.

One remaining possibility is if it is argued that it is *persons*, that is, beings actually capable of moral agency, who can go to heaven, and it is then stipulated that, by definition, a person only comes into existence with the first SFA. This would solve the problem (though we still have the “thin character” problem discussed earlier if death occurs after only one or two SFAs), but at the cost of denying that babies or even young children will have lives beyond the grave, because they would be non-persons. This, I think, is too steep a price for the vast majority of theists to pay.

## ii) Purgatory

Jerry Walls has argued for a conception of purgatory not as a place of punishment, but of post-mortem sanctification, to perfect us for heaven. He sees the necessity for this to follow from two “basic facts:” first, IMPECCABILITY, and second:

The great majority of persons—all, according to many theological traditions—are far from perfect when they die... [T]his is true even on the assumption that everyone has made at least some progress in the pursuit of holiness, some more than others. The

obvious fact remains that most are not completely holy, let alone impeccable, when they die.<sup>54</sup>

These two theistic commitments suggest that there must be a post-mortem sanctification process lest heaven be entirely unpopulated, and they combine with a third, that “we [humans] must cooperate in our sanctification”<sup>55</sup> to suggest that that post-mortem process should not simply be God “zapping” us into perfection, but rather a more drawn-out process, which requires a non-earthly, non-heavenly location to take place in.

Walls’s intended role for purgatory is thus as a kind of finishing school for the character. Will it be suited for tiny students? If so, who is to be their teacher? Do we acquire instant knowledge of the one true faith upon entering purgatory, or is it just like Earth, only one exits at sanctification (or, presumably, proving oneself irredeemable, in which case a trap door opens) because there is no exit via death? Its denizens (purgatorians?) must be capable of SFAs for it to serve its purpose, so presumably, on the Kane-Sennett view, evil must also occur there. How is it a better place to become perfect than Earth, then? Is it just that we have an infinity of time to work on it?

Let us set aside philosophical and theological objections to the very idea of purgatory<sup>56</sup> and focus on a more pressing problem. While purgatory might offer a way to address the Little Angel Automaton problem, it does so only at the cost of critically exacerbating APE. If newborn babies (or possibly even embryos, depending on the cut-off line for personhood and thus potential redemption) who have yet to have had the chance to make a single decision before dying and exiting Earth can, in purgatory, mature to fruition and perfect their souls while simultaneously performing the undetermined SFAs that allow for freedom in heaven, then we *really* have no need of Earth. Previously we could at least



claim there was a point to Earth as a place for the indeterminism necessary for heaven from which it was banished, but with purgatory, that role is taken.<sup>57</sup>

### iii) Two Kinds of Value, Objective and Subjective

One final, radical response to the problem of Little Angel Automatons is to bite the bullet and agree that yes, they would not have freedom, but assert that this would not matter *to them*. They would be just as happy as those who do. There is some plausibility in this: it makes sense to say that Eartha and Celeste are having exactly the same subjective experience of heaven, and indeed, feel identical beatific joy.

What does this say about the value of freedom? Sennett wanted to argue that the redeemed were free in heaven, presumably to support SUBLIMITY, the theistic claim that heaven is a better plane of existence for the redeemed than their life in heaven, and given TVF, the claim that freedom is necessary for something of transcendent value. But this move claims that some of the redeemed can be equally happy without the capacity for freedom (in fact, given that they died before a single SFA, they never even had freedom on Earth). How can this be sustained?

The only possible way would be to suggest that somehow heaven is *subjectively* better, but that *objectively* it is better that there be an Earth. That is, the transcendent value of freedom is an objective one: it is just a fact that a universe that contains, say, truly freely-given love, or “morally significant” freedom, which requires undetermined choice where evil is an option (to give two favorite candidates of free will defenders), is better than one without such things, but that this value is *not felt* by any of the inhabitants of the universe. What they feel, and what they *value*, are simply feelings, which may or may not be triggered by the exercise of genuine freedom.

This tack would definitely be novel. Too novel, I think, for any but the most radical theist to endorse. It also raises all kinds of questions: how many free beings must a universe contain to achieve the objective value? Is it quantifiable, so that more free beings means greater value? But then why are so many potential free beings allowed to die without exercising a single SFA? And it also raises yet another, new, variant of the problem of evil (and APE). If the exercise of true freedom is of no value to *us*, why is it that we have to suffer to produce it? There would certainly be more *subjective* happiness in a universe that contained no Earth (or purgatory) but only a heaven stuffed with Celeste-like beings (hence APE again). If God created the one we have over that one, he has favored objective value over the happiness of his creatures. God is using people on Earth (and requiring us to suffer all the evils that are present here) for some purpose that has nothing to do with any good that we could experience. We are all just pawns in some greater game. It cannot be said that we experience any benefit from the freedom we get on Earth, because the freedom we have in heaven is, if anything, better for us, and worse, the *same* for us as the faux freedom of Celeste and the Angel Automatons is for them.

### 13. What's a Theist to Do?

I have argued that it is impossible to reconcile the view of freedom needed for theistic responses to the problem of evil with a belief that the redeemed in heaven can act freely without the risk of sin. James Sennett's attempt at a reconciliation with a two-level conception of freedom fails for a variety of reasons. It rests on Robert Kane's work that is not just unconvincing but unable to provide the kind of freedom on Earth required by the free will defender to absolve God of responsibility for evil. Even if Sennett could provide a different basis for his two-level view than Kane's, however, his view would either consign

millions of very young infants to an eternity of unfreedom, or rely on there being a purgatory whose existence renders our Earth's at best redundant and at worst an unnecessary risk to our eternal fate.

A theist convinced by my arguments faces a choice of beliefs to discard: belief in heavenly freedom, belief in a heaven free from evil, or commitment to the use of free will in a defense or theodicy. It's not a happy choice. Self-interest would seem to lie with abandoning the last, in the hope that we can come up with an alternative explanation of the evil we see all around us, because, after all, the earthbound section of our existence is supposed to be an infinitesimal fragment of the whole. Heaven is too important a part of theism to compromise.<sup>58</sup>

#### References

- Baker, Lynne Rudder. 2003. "Why Christians should not be Libertarians: an Augustian challenge." *Faith and Philosophy* 20(4): 460–417.
- Dennett, Daniel. 1984. *Elbow Room*. Cambridge, MA: MIT Press.
- Frankfurt, Harry. 1969. "Alternate possibilities and moral responsibility." *Journal of Philosophy* 66 (3): 829–39.
- Frankfurt, Harry. 1971. "Freedom of the Will and the Concept of a Person." *Journal of Philosophy* 68 (1): 5–20.
- Hick, John. 1966. *Evil and the God of Love*. New York: Palgrave Macmillan.
- Hume, David. 1988. *An Enquiry Concerning Human Understanding*. Buffalo: Prometheus.
- Kane, Robert. 1996. *The Significance of Free Will*. Oxford: Oxford University Press.
- . 1999. "Responsibility, Luck, and Chance: Reflections on free will and indeterminism." *Journal of Philosophy* 96 (5): 217–240.

- . 2005. *A Contemporary Introduction to Free Will*. Oxford: Oxford University Press.
- Lewis, David. 1981. “Are We Free to Break the Laws?” *Theoria* 47 (3): 113–121.
- Mackie, J.L. 1955. “Evil and Omnipotence.” *Mind* 64 (254): 200–212.
- . 1982. *The Miracle of Theism*. Oxford: Oxford University Press.
- Parfit, Derek. 1995. “The unimportance of identity.” In *Identity*, edited by Henry Harris, 13–45. Oxford: Oxford University Press.
- Pawl, Timothy and Kevin Timpe. 2009. “Incompatibilism, Sin, and Free Will in Heaven.” *Faith and Philosophy* 26 (4): 396–417.
- Plantinga, Alvin. 1977. *God, Freedom, and Evil*. Grand Rapids: Eerdmans.
- . 2009. “Transworld Depravity, Transworld Sanctity, & Uncooperative Essences.” *Philosophy and Phenomenological Research* 78 (1): 178–191.
- Sennett, James F. 1999. “Is There Freedom in Heaven?” *Faith and Philosophy* 16 (1): 69–82.
- Swinburne, Richard. 1996. *Is There a God?* Oxford: Oxford University Press.
- Van Inwagen, Peter. 1984. *An Essay on Free Will*. Oxford: Clarendon.
- . 2008. “How to Think about the Problem of Free Will.” *Journal of Ethics* 12 (3/4): 337–341.
- Walls, Jerry. 2015. *Heaven, Hell, and Purgatory: A Protestant View of the Cosmic Drama*. Grand Rapids: Brazos Press.
- Zagzebski, Linda. 1991. *The Dilemma of Freedom and Foreknowledge*. New York: Oxford University Press.

## Notes

1. I include in this not just “free will defenses” but also John Hick’s “soul-making” theodicy—see note 31.

2. Strictly speaking, a “defense” against the problem of evil is usually taken to mean a demonstration that it has not been proven that an omnipotent, all-loving God is inconsistent with the existence of evil. This is weaker than a “theodicy,” whose claims are asserted as being actually or probably true rather than merely not proved false. I will use the term loosely (as has become common) to cover both defenses and theodicies, and the term “free will defender” to refer to those who are committed to either.
3. Sennett 1999.
4. Strictly speaking the problem can be expressed without this claim, but the claim that God created the universe the way it is, and could presumably have made it otherwise, certainly makes the problem vivid.
5. For those steeped in the work of Alvin Plantinga (see, e.g., his 1977, 2009) or other molinists, the phrase “freedom is such that...” can be read as “it is possible that the counterfactuals of freedom are such that...”
6. The nature of the phenomenon depends on the version of the defense/theodicy. Popular candidates include actions of moral worth, freely given love, and the mere exercise of genuine freedom.
7. See for example Plantinga 1977, part I and Swinburne 1996, chapter 6.
8. There are views that assert the existence of a heaven that is as yet empty of (ex) humans. (See John 3:13, but contrast 2 Kings 2:11.) For such views claims like this one and the others about heaven can be re-parsed into the future tense
9. “To use a classic theological term, those in heaven must be fully perfect in character in such a way that they are ‘impeccable,’ which means that they can no longer sin. Doing evil must be impossible for the redeemed in heaven.” Walls 2015, 94.

10. “Best” is vague here. Pawl and Timpe offer: “heaven is essentially a place of ultimate happiness” and “the state of human existence than which none more desirable can be conceived,” Pawl and Timpe 2009, 401.
11. Sennett 1999, 79, note 4.
12. Remember that the free will assumed by free will defenders has to explain all the horrors that we see on Earth. How would heaven be better if its inhabitants are capable of the same atrocities? Even if the worst offenders would be in hell, IEF seems to require that evil be done, and good people can turn evil if evil is done to them. And suppose heavenly inhabitants do commit evil acts. Should they be thrown out *a la* Lucifer? This causes eschatological problems: can hellions also get promoted for good behavior? Is there a revolving door in the afterlife? Or does heaven steadily empty? Alternatively, if sinners get away with it in heaven, that seems unfair on those punished for earthly sins. Why should the *location* of your sin’s commission rather than the nature of the sin itself determine your punishment?
13. It used to be that Catholic Doctrine condemned the unbaptized to hell, which would certainly preclude piously-and-altruistically-motivated abortions, but this doctrine has softened steadily, first to the suggestion that those infants would go to limbo, and then even limbo was phased out in 2007 with the publication of “The Hope of Salvation for Infants Who Die Without Being Baptized” ([http://www.vatican.va/roman\\_curia/congregations/cfaith/cti\\_documents/rc\\_con\\_cfaith\\_doc\\_20070419\\_un-baptised-infants\\_en.html](http://www.vatican.va/roman_curia/congregations/cfaith/cti_documents/rc_con_cfaith_doc_20070419_un-baptised-infants_en.html)). For more on limbo, see chapter 12.
14. The view is so named because it is the view that freedom is compatible with (causal) *determinism*, which is the view that every event that happens does so in a way that is entirely predictable given the combination of background conditions and laws. So, for example, given the state and laws of the universe at any particular moment, there is only

one action it is physically possible for any individual to perform: that person is *determined* to do so. (This is not to say that is *logically* impossible that they do otherwise, as compatibilists often point out: see for example Lewis 1981.)

15. There are exceptions—see for example Baker 2003. But even she begins by quoting Zagzebski to the effect that the belief “that human beings have free will in a sense...that is incompatible with determinism” is “central to Christian practice.” (Zagzebski 1991, 3.)
16. Kane 2005, 149. See also Pawl and Timpe 2009, 399.
17. Van Inwagen 1984, 16. Cited by Sennett 1999, 72.
18. There is controversy over where to draw the border line between internal and external, which we can gloss over here. We will also treat “self” and “character” as interchangeable for the time being, but look into a possible distinction later.
19. Possible worlds are favorite tool for philosophers. Roughly a “possible world” is like the parallel universes beloved of science fiction. It is not part of our universe, it is an alternative to it. So another possible world where I am president is an entire alternative universe that contains somebody whom we can (for reasons that differ among philosophers) identify as me, except that he is president. Another possible world cannot be reached from this one and philosophers differ on whether they are real or not. But they are useful ways of thinking of “counterfactuals”—that is, ways things could have been even if they actually weren’t. The online entry in the Stanford Encyclopedia of Philosophy is, as always, a great resource for those unfamiliar with the topic.
20. Imagine there was a particularly savage serial killer (the Deranged Disemboweler) on the loose in your city. Hearing of this killer’s gruesome crimes, your mother calls you up and asks, in all earnestness, “Are you the Deranged Disemboweler?” When you respond

in outrage, she seems surprised and says, “But don’t you want true freedom? And that requires that all options be open to you, no matter what your character.”

21. Hume 1988, 90-91.
22. Kane 1999, 237.
23. Dennett 1984, 133, Kane 1996, 38, 2005, 81. Sennett 1999, 73 has “I cannot do otherwise,” which substitutes clarity for euphony.
24. “Possible” in the sense of “possible for Luther in the actual world with its actual laws.” Both compatibilists and libertarians can agree that other actions are *logically* possible.
25. J.L. Mackie famously argued (Mackie 1955, 1982) that even were we to be free in a libertarian sense this would not be sufficient to make IEF plausible. I am inclined to agree with him, but many have argued otherwise, and I will set that debate aside in this paper.
26. Why should I be praised for helping an old lady across the road if I could equally well have pushed her under the bus, given exactly the same character? Why should I be blamed for murder if my character does not determine it and there’s a possible world where, with exactly the same character, I lead a blameless life?
27. The phrase itself is ambiguous, because the “self” part could be the active *former* or the passive *formee*. However, it seems clear that Kane intends it to be *both*. This sounds a bit like pulling oneself up by one’s own bootstraps, and we will devote a later section to scrutinizing Kane’s conception in detail.
28. Kane 2005, 130-1.
29. Sennett 1999, 74.
30. Sennett 1999, 75-6.
31. An anonymous reviewer suggested to me that one might simply deny that their characters *were* “identical” because one’s character includes moral virtues, and whether



or not one has moral virtues depends on the history of their formation. This is an idea put to use by John Hick's "soul-making" theodicy (Hick 1966), which plainly addresses APE by suggesting the point of Earth is so that humans can have a hand in creating their souls by making free choices. The literature on Hick's theodicy is voluminous (troubling apparent implications of it is are that it appears to place a limit on God's power, and that humans who have gone through "soul making" have a kind of virtue unavailable to the angels and even to God), so I will not directly challenge it here (although I have more to say about it in chapter 12), except to note that its reliance on free will opens it up to problems I raise below, most notably the "little angel automatons," and that if purgatory is a possible site for soul-making then APE recurs. As to the idea that whether or not one's current action is virtuous depends on the history of one's character formation, I address what I take to be the beliefs that might motivated it in what follows, especially the android example.

32. The topic of personal identity, particularly over time (that is, settling the issue of whether, and if so, in virtue of what, *current-me* is *the same person* as *me-aged-ten*) is a much-discussed topic in philosophy. The chapters by Manninen, Olson, and Guillon in this volume provide good examples. Derek Parfit has done more than any other philosopher to popularize the use of Star Trek transporters in such discussions—for a fairly accessible example, see Parfit 1995.
33. Notice that this would include qualities that are inessential to the issue of personal identity through time. That is, I used to hate beets but now I don't, and yet I am the same person throughout. Nonetheless, my early avoidance of beets can be explained by my childhood "thick" self, which included a beet-aversion, while my current willingness to eat beets can be explained by my adult beet-acceptance.

34. Perhaps you might make a case that my later beet-acceptance was something I chose if I made a conscious decision to overcome my distaste for them and forced myself to eat them over a sustained period until I had beaten my former revulsion into submission. But, even if so, that wouldn't change the fact that the early dislike was something I acquired involuntarily from forces (genetics, environment) beyond my control.
35. The dispute over whether or not one's sexuality, for example, is "chosen" or innate could be couched in terms of whether or not it is a product of (or amenable to alteration by) self-formation.
36. Kane (1996, 2005) is a great resource for the philosophical literature on this point. But such character manipulation is also a trope in a lot of fiction. The books and films *The Manchurian Candidate*, *A Clockwork Orange*, and *Total Recall* (based on the Philip K. Dick short story "We Can Remember It for You Wholesale") suggest ways this might work. Your would-be controllers could manipulate your desires so you want to kill people they want assassinated, or you can no longer do the violent acts you used to enjoy without being violently ill, or your memories could be altered so that you *believe* you are a person whose life fits their needs.
37. Sennett 1999, 74-5.
38. Mill, an avowed compatibilist, asserted, "we are exactly as capable of making our own character, *if we will*, as others are of making it for us" (J.S. Mill, *A System of Logic*, Book 6, Chapter 2, §3.) and in fact there are several compatibilist accounts of self-formation. Harry Frankfurt (1971) suggested (very roughly) the following. One's character can contain conflicting desires: the desire to have a cigarette and the desire to stop smoking, for example. Desires like the second ("second order" desires—desires *about* one's own desires) are more integral to one's true character, and it is in allegiance to those that one can alter the lower-level aspects of one's own character by, for

example, isolating oneself from cigarettes until the urge (ideally) disappears. Surely these are cases where the self both is in charge of change and is itself changed?

However, Frankfurt's view has been challenged by many critics—again, Kane provides a good overview (1996, 2005).

39. “[T]he standard compatibilist doctrine of free actions as those determined by the agent’s character escapes the Consequence Argument only if the agent’s character was formed, at least in part, by proximately undetermined free actions by the same agent,” Sennett 1999, 75.
40. Sennett 1999, 75.
41. All quotes that follow are from Kane 1999.
42. Kane 1999, 225.
43. Or becoming the Deranged Disemboweler (see note 20 above).
44. “There is a tension and uncertainty in our minds at such times of inner conflict... a kind of stirring up of chaos in the brain that makes it sensitive to microindeterminacies at the neuronal level. As a result, the uncertainty and inner tension we feel at such soul-searching moments of self-formation is reflected in the indeterminacy of our neural processes themselves,” Kane 1999, 224-5.
45. Consider Eve. Why is she wrong to eat the apple? Because God told her not to. But if her character was perfect to begin with, then not only would she never consider disobeying him, she would never indulge in SFAs that would form her character into one that would consider disobeying him, and the Fall would have been averted. And that’s not even to get into the character of the Serpent.
46. Pawl and Timpe 2009, 408: “If both of these actions are consistent with the nature of heaven and one sees good reasons for engaging in both activities, then one’s moral

character needn't determine one's choice either way. ...[T]here can be non-derivative free choices even in heaven."

47. Perhaps, instead of being "zapped" into perfection (to use Jerry Walls' terminology—Wall 2015, 94, 112) Jezebel could take a trip to purgatory and continue working on her self until she has finally reached Mary's level. We will discuss purgatory below, but notice that (a) Jezebel has to work a lot harder than Mary through no fault of her own, and may in fact never make it, and (b) this wouldn't help Rock, who cannot change.
48. Strictly speaking, if we make this terminological switch we should be discussing CFAs rather than SFAs in what follows, but to avoid yet more acronyms we'll retain the former.
49. "A Crusade Born of a Suffering Infant's Cry," *New York Times*, March 19, 2005.
50. Consider Plantinga's suggestion for the source of natural evil: "Satan, so the traditional doctrine goes, is a mighty non-human spirit, who, along with many other angels, was created long before God created man. Unlike most of his colleagues, Satan rebelled against God and has since been wreaking whatever havoc he can. The result is natural evil," Plantinga 1974, 192. We are apparently asked to believe that, to focus solely on the case of Hallopeau-Siemens syndrome, it is more important that Satan be free to torture babies (and in a different way, their parents and caregivers) than that God step in and undermine some of his efforts
51. They might even be *incredibly* rare: how often would we have *exactly* equal strengths of reasons for both sides of a decision?
52. The number goes up precipitously if you count the numbers of fertilized eggs (possibly up to a half of all conceptions) that modern medicine tells us spontaneously abort, often without the prospective mother even knowing they were fertilized.
53. Kane 1999, 224.

54. Walls 2015. 95.
55. Walls 2015, 113. This should remind us again of Hick's "soul making" (see note 31 in section 0), which suggests that the problem described below is also a problem for that theodicy.
56. Positing an intermediate state to bridge the gap between heaven and Earth might strike a reader who until now had not considered that there might be a third post-mortem state as *ad hoc* and ontologically lavish. That said, it is an idea familiar to Catholics in particular. But this fact also reveals that it is *theologically* divisive. Walls himself says that "*purgatory* is a fighting word" and cites Calvin's full-throated rejection of the notion: "purgatory is a deadly fiction of Satan, which nullifies the cross of Christ, inflicts unbearable contempt upon God's mercy, and overturns and destroys our faith," Walls 2015, 91-2.
57. Or to put it another way: Walls's purgatory is a place where one may perfect one's character ready for heaven. If so, then one of two things must be true: either there is no evil possible in purgatory, in which case IEF is undermined, or it is just like Earth, with the possibility of evil, except that you cannot die, in which case what is the point of Earth? Earth would then be rather like a terrifying hoax, deceiving us about the necessities of our existence, possibly frightening us into actions that could condemn us into damnation down the line.
58. I would like to thank an anonymous reviewer for helpful comments on an earlier version of this chapter.