Religious Studies (2014) **50**, 255–269 © Cambridge University Press 2013
doi:10.1017/S0034412513000401

Augustine, the origin of evil, and the mystery of free will

ADAM M. WILLOWS

Department of Theology and Religion, University of Durham, Abbey House, Palace Green, Durham, DH1 3RS, UK
e-mail: adam.willows@durham.ac.uk

Abstract: The question of why humanity first chose to sin is an extension to the problem of evil to which the free-will defence does not easily apply. In *De libero arbitrio* and elsewhere Augustine argues that as an instance of evil, the fall is necessarily inexplicable. In this article, I identify the problems with this response and attempt to construct an alternative based on Peter van Inwagen's free will 'mysterianism'. I will argue that the origin of evil is inexplicable not because it is an instance of evil, but because it is an instance of free will.

Introduction

In this article, I defend and refine Augustine's response to the problem of evil with reference to contemporary thought regarding free will in the philosophy of mind. My primary intent is not exegetical; rather, I aim to use Augustine and others to develop an interesting and defensible philosophical position. I will begin with a short description of the problem of evil and the free-will defence before looking at Augustine's account. Augustine identifies a further problem – namely how to account for the origin of evil. One way in which he responds to this is to claim that the origin of evil is fundamentally inexplicable. I will follow Robert Brown in arguing that this approach is the best of Augustine's responses. However, I will show that there are different ways in which this argument can be made, and that Augustine's position can be strengthened. Augustine's approach is to connect inexplicability to the nature of evil. I will look at the problems with this approach, and suggest that the inexplicability should be located in all human will, whether or not it is evil. In order to show this, I will argue for Peter van Inwagen's claim that all the available accounts of free will are flawed, and so free will must be a mystery.

I hope to show that this part of the existing discussion on free will can be helpful for the claim that evil is inexplicable.

Free-will theodicy and scope

The problem of evil is the claim that the existence of evil is incompatible with a God who is both omnipotent and wholly good. Since evil does exist, such a God cannot exist. It appears that the theologian must adhere to the three premises 'evil exists', 'God is omnipotent', and 'God is wholly good'. However, given the further clauses that good will always attempt to eliminate evil, and that there are no limits to an omnipotent being's powers, it seems that a God who was both wholly good and omnipotent would eliminate evil completely. The task for the theologian is to explain how the three premises are compatible (Mackie (1990), 25–26).

The free-will defence is one such explanation. It argues that it is better for there to be creatures with free will and the consequent ability to do both good and evil, than for such creatures not to exist. God, being wholly good, has created humans with free will and we are responsible for evil by the exercise of our free will. If God were to stop us doing evil we would not be free, and therefore would be unable to do good. It is better that God allows us to be free, even though our freedom has resulted in evil. Therefore God's goodness and omnipotence are compatible with the existence of evil (Evans (1999)).

The origin of evil and responsibility

The free-will defence is used by Augustine to respond to the problem of evil: 'So too free will, without which no one can live rightly, must be a God-given good, and you must admit rather that those who use this good wrongly are to be condemned than that He who gave it ought not to have given it' (Augustine (1955), 131). However, Augustine notices a further problem, which the free-will defence does not immediately resolve (Matthews (2004), para. 5). It is logically possible for a being to have free will and not sin at all. Why then did the first humans choose to sin? It cannot be through some defect in their character or will, since God must have created them wholly good. If he did not, then he would be responsible for their flawed nature and thus responsible for the fall – the origin of evil. Nor can it be that they were compelled to do so, since only something superior to the will can compel it – but anything superior would be even more wise and good and so would not compel it to do evil (Evans (1982), 115). Although the free-will defence may show that God's existence is compatible with the existence of evil, it does not show that God's existence is compatible with the origin of evil.

Augustine attempts to answer this problem by claiming that the origin of evil (the first instance of evil willing) is necessarily incomprehensible. He uses the idea

of evil as privation – that is, the absence of good (Mann (2001), 40–48). If evil is simply an absence or a nothing, then it has no true existence and hence cannot be known: 'sin . . . is a defective movement, and a defect comes from nothing' (Augustine (1955), 137). Augustine is relying on the idea that to comprehend something is to understand its cause. The first evil will has no cause, and so it cannot be understood. Chappell calls this the NEA – 'No-Explanation Account' (Chappell (1994), 869–884).

In what follows, I intend to refine and defend the NEA. Before I proceed, I want to make two preliminary points. The first is to do with the relationship of the NEA to theodicy and the aim of this article. The NEA is Augustine's explanation of the origin of evil, which he has correctly identified as a problem for the free-will defence. Any theodicy based on the free-will defence will need to answer this problem in some way. Such an answer, however, can only represent part of a complete theodicy. Supporters of the free-will defence must still address several other criticisms. These include the argument that they cannot account for evils which do not seem to be related to human actions ('natural' evils) and the claim that they do not do enough to address particularly terrible evils (Pereboom (2004), 149–155). These create important questions for any theodicy. However, my goal in this article is not to provide a complete theodicy, but to resolve the specific problem of the origin of evil. Although the other challenges to a free-will theodicy are important, they are not within the scope of this article. If Augustine's NEA (or my revised version) is successful, it may form part of a successful free-will theodicy and for this reason it merits investigation.

My second point concerns moral responsibility. My aim here is to show why it is important for Augustine to break the causal link between God and the fall. Augustine uses the NEA because it breaks this causal link; if the NEA is correct the flaw in creation does not occur as a direct causal result of God's actions. The assumption is that if God is not causally linked to the fall, then he is not necessarily to blame for it. Strictly speaking responsibility has to do with control, rather than causation (Duff (1998), para. 4). In the scenarios I am dealing with this does not particularly matter; in every case addressed here in which an agent is the cause of a situation, they are also in control of it. For the remainder of this article, I will assume that causation is a sufficient condition for responsibility; if an agent is directly causally responsible for something they are morally responsible and potentially morally to blame for it. The NEA aims to show that God is not responsible for the fall because he is not the direct cause of it. Showing that God is not causally responsible for the fall would not necessarily absolve him of moral responsibility. There are other ways in which an agent might be morally responsible for something, and I discuss two ways in which God could be considered responsible below. What a successful NEA would do is show that God is not *necessarily* morally responsible for the fall. If the direct causal

https://doi.org/10.1017/S0034412513000401 Published online by Cambridge University Press

connection is preserved, God must be morally responsible; with it removed, Augustine is free to argue that other agents alone are morally responsible, as in fact he does.

Before moving on I want to dismiss two other ways in which God might be deemed responsible for the fall without being the direct cause of it. The first is that God is causally linked to every event through primary causation. Although he might be deemed responsible in some sense as a result of this link, he is not morally blameworthy. As the free-will defence points out, it is better that God creates in the way that he does than not, despite the fact that this allows us to do evil. The second way is that even if God is not the direct cause of evil, he might be blamed for taking too much of a risk, creating a scenario in which evil is likely to arise by creating beings who can fall. In practical terms, it is impossible for us to analyse the risk involved here, although given divine omniscience it would presumably be possible for God. The risk might be extreme (if the first humans have infinite time in which to fall) or minimal (it is difficult to see why non-fallen beings would choose to fall). In any case, probability alone is not the best way to analyse acceptable risk. There are some scenarios in which a probability of 1/10 of an event occurring is an acceptable risk; others in which it is not. A better analysis is suggested by Sven Hansson: 'Exposure of a person to a risk is acceptable if and only if this exposure is part of an equitable social system of risk-taking that works to her advantage' (Hansson (2003), 305). I suggest that under this analysis, even if God did take a risk in giving us free will the risk was acceptable since free will is such a large benefit. Only if he is directly causally responsible for the origin of evil, then, will God be culpable. It is this causal responsibility that the NEA denies. However, it is not Augustine's only attempt at explaining the origin of evil. In the next section, I will examine his other answers and the problems with them to show that the NEA is the best way to proceed.

Alternative explanations of the fall

Robert Brown identifies three other explanations of the fall in Augustine. First, that free creatures have an inherent weakness; second, that pride is the cause of the fall; and third, that we cannot understand the cause of the fall although it is comprehensible to higher beings (Brown (1978), 317). Augustine's first explanation is that all things (including humans) are 'subject to change, because they were made not out of his being but out of nothing' (Augustine (2003), 472). According to Brown, this mutability and creation out of nothingness means that humans necessarily fell. Chappell disagrees with Brown here. He argues that Augustine sees humanity's creation out of nothingness as a necessary but insufficient condition for the fall – that is, only creatures created out of nothingness can fall, but it does not follow that they *must* fall (Chappell (1994), 875–877). If Chappell is correct then Augustine is not offering an explanation

https://doi.org/10.1017/S0034412513000401 Published online by Cambridge University Press

for the fall, but throwing light on the conditions for it. I shall follow Brown's interpretation to show that even if Augustine is attempting to explain the fall here, the explanation is not a satisfactory one.

If Brown is correct then Augustine believes that evil has its origin in humanity's changeable nature. There are two ways in which this view could be interpreted, both of which cause serious problems. The first is that God chose to create humans out of nothingness. If this is the case, then it is God who is responsible for the fall by not creating humans with a wholly good nature, leading them to sin. The second possibility is to suppose that God cannot but create out of nothing. This approach absolves God of responsibility for the fall, but it seems inconsistent with Augustine's belief that there are free beings other than God who have not sinned (the good angels). Ultimately, the problem with this explanation of the fall is that on either interpretation, humans are not responsible for their sin, since they are determined by their nature and do not sin freely. Augustine acknowledges that 'punishment and reward would be unjust, if man did not have free will' (Augustine (1955), 76). Because of this, his account of humanity as created out of nothingness does not provide an explanation for the fall.

The second explanation of the fall is that it is a result of pride: 'could anything but pride have been the start of the evil will?' (Augustine (2003), 571). Pride is the turning of the will towards what the self desires rather than what God desires. This argument has an obvious flaw. Any evil will or act caused by pride could not be the origin of evil, since at least one evil – pride – would already exist. If pride comes before the first evil will, then it is not the cause of the fall – it *is* the fall. Rather than being 'what was the cause of the first evil will?' the question becomes 'what was the cause of pride?'. This is well put by Brown: 'Pointing to pride . . . is only the substitution of a synonym for the inexplicable free act of falling' (Brown (1978), 322).

The third alternative that Augustine offers is closer to the NEA. He says that although there is a cause of the first sin, we cannot understand it in our current state. Unfortunately, here Augustine retains the unsatisfying aspects of the NEA while discarding its strengths. There is nothing necessarily wrong with an explanation of the fall, and it is not useful to avoid providing one unless it serves some purpose. The reason that the NEA claims that there can be no explanation for the fall is that it seems that any explanation must involve identifying a cause of the fall, and that any such cause must have been created by God, making him responsible. By claiming that there is a cause that we cannot understand, this argument sustains the link between God and the fall which the NEA breaks. Furthermore, Brown points out that Augustine speculates elsewhere on various aspects of humanity's condition prior to the fall (*ibid.*, 323). This weakens his attempt to claim that the pre-fall state of the human will is unknowable, if other areas of human nature are knowable.

https://doi.org/10.1017/S0034412513000401 Published online by Cambridge University Press

Problems with the NEA

I have shown that Augustine's other attempts to account for the origin of evil do not succeed. I will now return to the NEA, and potential problems with it. To reiterate, the NEA is Augustine's claim that the origin of evil is without a cause and so fundamentally inexplicable. If Augustine is correct, then this breaks the causal link between God and the fall – it allows that free creatures may sin without God being responsible for the aspect of their character that causes them to do so. There are two problems for Augustine's version of the NEA. I think that one can be resolved, but the other reveals an inconsistency in Augustine. I will look at these two problems before attempting to reformulate the NEA to make it stronger.

The first problem involves a dilemma proposed by Augustine himself: 'If the first man was created wise, how was he seduced? If he was created foolish, why is not God the cause of vice?' (Augustine (1955), 213). Augustine's response is to claim that there is a middle transitional state between wisdom and folly which allows man to be seduced without the blame falling on God. However, Chappell offers a reformulation of this dilemma which Augustine's response does not resolve and to which, he claims, the NEA is vulnerable. This revised dilemma is based on the question 'Did God create Adam morally perfect or morally imperfect?' (Chappell (1994), 873). Its four parts are as follows:

1: *Premise*: God created humans, but humans are responsible for the fall.
2: Humans were created either (A) morally perfect or (B) morally imperfect.
3: If (A) the fall was an impossibility, since a morally perfect being cannot fall. Therefore 1 is false.
4: If (B) then the fall was God's fault for creating humans morally imperfect. Therefore 1 is false.

Chappell argues that the NEA cannot answer this criticism. He says that the NEA's response ought to be to attack 3 by saying that although the fall is impossible, it still happened (hence the inexplicability). He rightly says that this is not a good response, since the impossibility referred to in 3 is a logical impossibility based on the definition of a morally perfect being. The inexplicability in the NEA instead refers to a practical inexplicability or impossibility, rather than a logical one. The only other possibility he offers is to attempt a redefinition of 3's morally perfect person. Chappell argues that any such redefinition to allow a morally perfect being to fall will run a serious risk of collapsing moral perfection into moral imperfection. He concludes that the NEA alone is unable to answer this dilemma, and goes on to offer a response based on his interpretation of Augustine on our creation from nothingness discussed above.

I think that Chappell is mistaken here. There is a flaw in the dilemma which allows the NEA to respond. The weak point is not 3, but 2, which depending upon

the definition of moral imperfection is either false or ineffective. There are two ways in which moral imperfection could be understood here. First, it could be the inverse of moral perfection as defined in 3; that is, a morally imperfect being is one which must necessarily fall. Call a premise using this definition 2i:

2i: Humans were created either (Ai) unable to fall or (Bi) determined to fall.

Both 3 and 4 follow from 2i. However, 2i is false. As previously mentioned it is logically possible for a being to be free but never in fact to sin. Such a being would be neither unable to fall (Ai) or determined to fall (Bi). It is clear from Augustine's discussion on free will that he sees humans as exactly this kind of being – free will is identified as a good, but also that by which we sin. On this definition of moral imperfection, then, there is a middle ground between the horns of the dilemma.

The second possible definition of moral imperfection does not allow for such a middle ground. On this definition, a morally imperfect being is one which is able to fall. This includes the first definition of moral imperfection (determined to fall) as well as creatures with free will such as humans, who can fall but need not. Call a premise using this definition 2ii:

2ii: Humans were created either (Aii) unable to fall or (Bii) able to fall.

I think that 2ii is a more faithful reworking of Augustine's original dilemma. In this case, 2ii is correct and 3 follows from it. However, 4 does not. If God created humans able to fall but not determined to, then it is not his fault that they did. The further question then arises of why exactly humans fell, and it is this question which the NEA is designed to answer. It allows the response that Augustine wants – that God created humans sinless but with free will, and that they fell of their own accord.

I have shown that Chappell's criticism is ineffective. I now turn to another problem with the NEA, which I believe identifies the need for a reworking of the argument. The NEA requires a reason for the inexplicability of the origin of evil. Augustine's reason is that the first evil will comes from nothing, and that nothing is incomprehensible: 'that which is nothing cannot be known' (Augustine (1955), 136). Elsewhere, he says that the cause of the evil will is a defective cause – that is, the absence of good: 'To try to discover the causes of such defection – deficient, not efficient causes – is like trying to see darkness or hear silence' (Augustine (2003), 480). I do not think that Augustine is clear about what he means here. He seems to conflate two possibilities – that the cause of the first evil will is an example of causation by omission, or that the first evil will is entirely uncaused. There is a problem with both of these possibilities, which I shall now investigate.

Causation by omission occurs when something is caused by the absence of a particular thing or event – for example 'the plant wilted because I did not water it'. Here the absence or non-existence of the event 'my watering the plant' is taken to be the cause of its wilting. Similarly, Augustine may be saying that the absence

https://doi.org/10.1017/S0034412513000401 Published online by Cambridge University Press

of good at the first instance of evil willing is the cause of that evil willing. This view seems to be supported by his discussion of 'deficient' causes. There are two problems with this approach. First, it may re-establish a causal link between God and the origin of evil. This is the view that Leibniz takes in his early thought by arguing that God is the cause of all privations as well as all realities (Murray (2005), sec. 3 para. 4). The second is that it is doubtful whether causation by omission is an example of genuine causation at all. Phil Dowe outlines some of the arguments against the view that causation by omission is genuine causation, including the fact that omissions do not seem to be genuine events and the problem that it is hard to establish a spatial or temporal link between omissions and effects (Dowe (2001), 216–226).

However, it is possible that Augustine is trying to make exactly that point; omissions are not genuine causes. This brings us to the second possibility hinted at by Augustine – that the first evil will is entirely uncaused. This is supported by his claim that the defect comes 'from nothing' (Augustine (1955), 137). Unfortunately, there is a problem with this view as well. If the first evil will is entirely uncaused, in what way can it be said to belong the agent? Surely for the first humans to have responsibility for it, they must be causally connected to it. If the origin of evil was a random occurrence, then it is not the fault of humanity.

It might be thought that a non-causal view of what constitutes human action could help Augustine here. Most action theorists think that what makes an event an action has to do with how it is caused; that is, an action is an event with a particular *kind* of cause (usually the cause is the agent's intention). Although it is not widely accepted, there is an alternative view. Thinkers like Elizabeth Anscombe and Harry Frankfurt argue that an action is connected to an agent by its goal or *telos*, rather than by its cause. They think that what makes an action 'belong' to me is that I direct or govern the event (Frankfurt (1978), 159). For example, the difference between my raising my hand and a spasm causing me to raise it is that in the first instance I am directing the hand-raising, whereas in the second I am not. They do not deny that intention plays a role in making an event an action, but deny that an intention functions as a cause in the normal sense. This kind of action theory seems to show how an agent could be responsible for an uncaused evil willing. If they direct or guide the event, then it 'belongs' to them. Unfortunately, it does not do the work that Augustine needs. Non-causal action theories do not deny that actions are caused; they deny that the cause is what makes the event an action (as opposed to a mere random occurrence). These theories do not allow for the kind of uncaused action that Augustine needs. A standard action theory argues that an action-event is caused by the intention. This kind of non-causal action theory argues that an action-event has both a cause and an intention distinct from the cause which makes the event into an action. They are called non-causal theories because they deny that the cause is what makes the event an action – not because they deny that actions have a cause. They still have

causes and hence can still be connected to God. Nor can Augustine plead that the first evil will is a special case – a unique example of an uncaused event. Such a position would be deeply unsatisfying. Although the event itself would be unexplained, Augustine would need to explain why that one event differed from every other event of evil willing. In any case, Augustine clearly thinks that every instance of evil willing comes from nothing, since he discusses our control over the defect in our will which comes from nothing (Augustine (1955), 137).

It is clear that both causal and non-causal accounts of the origin of evil cause difficulties. I suggest that for the NEA to work, Augustine needs to rely on a radical inexplicability which neither offers any kind of cause nor declares the impossibility of a cause for the origin of evil. Even his argument that the evil will comes from nothing comes too close to providing an explanation.

I want to mention another apparent reason for the inexplicability of the origin of evil. Theologians such as Terence Tilley, Kenneth Surin, and Karen Kilby have criticized theodicy. The specifics of these positions vary, but they all claim that providing an explanation is not the primary task of theological discussions of evil. Tilley in particular argues that theodicy is a mistake and should not be attempted at all. Instead theologians should identify, understand, and try to overcome evil (Tilley (2000), 250). It might be thought that this approach to evil could provide the inexplicability that Augustine is looking for. However, this is an error. Some of these approaches argue that evil *ought* to be unexplained, whereas Augustine's argument needs to show that evil *cannot* be explained, even if we want it to be. Although they may have similar conclusions, the claim that evil ought to be unexplained is not helpful for Augustine's argument.

The NEA reformulated

I think that there is an alternative way of showing why the origin of evil is inexplicable. Augustine focuses on the nature of evil to provide a reason for the inexplicability. I intend to focus instead on the nature of free will. I will defend Peter van Inwagen's free-will 'mysterianism' to show that the claim that all instances of willing are inexplicable is plausible. If this is indeed the case, then it provides a solution to Augustine's problems with the NEA. If free will is inexplicable then the fall, being an instance of (evil) willing, must also be inexplicable. The idea that free will is inexplicable may seem a surprising claim. However, given the intractability of much of the debate over free will, I believe that it merits investigation. This investigation is important because if van Inwagen is correct, then the NEA will become a plausible position regarding the origin of evil. In his essay van Inwagen assumes the existence of free will, although he makes an argument for its existence elsewhere (van Inwagen (1983), 23–54). As Augustine also makes this assumption, I shall do the same for the purposes of this article. My aim here is not to contribute further to the discussion on free will

https://doi.org/10.1017/S0034412513000401 Published online by Cambridge University Press

per se, but to show that a particular view of free will can be helpful for our understanding of the origin of evil.

There are two main views regarding the nature of free will. The first is compatibilism – the view that we can be determined (unable to do other than what we actually do) and still have free will. Compatibilists often claim that determinism is in fact necessary for free will. The second view is incompatibilism. This is the view that only by being undetermined can we have free will. Van Inwagen's argument is that although we undeniably have free will, both compatibilism and incompatibilism have serious problems with them. Van Inwagen concludes that free will must be a mystery, since 'free will undeniably exists and there is a strong and unanswered prima facie case for its impossibility' (van Inwagen (2002), 159). In what follows I will explain his arguments against both compatibilism and incompatibilism and defend them against critics.

The 'Consequence Argument' is a famous argument against compatibilism, and is van Inwagen's method of attack. He states that determinism is the claim that in any possible world any state of affairs is a result of the combination of the laws of nature and the state of the world at some time in the past. He then points out that no one has any choice about either the laws of nature or the past. Therefore, if determinism is true, no one has any choice about anything. If someone has no choice about anything, van Inwagen concludes, she is not free – so determinism is not compatible with free will.

In attempts to respond to the consequence argument, compatibilists have attacked van Inwagen's understanding of freedom. Van Inwagen sees freedom as the ability to do other than we actually do, known as the 'could have done otherwise' definition. The consequence argument shows that this definition is incompatible with determinism. Compatibilists have suggested that freedom does not have to do with the ability to do otherwise. Instead they say that it involves our decisions having an effect on our actions, even if our decisions are determined. Daniel Dennett suggests that this is what we 'want' from free will (Dennett (1984), 138–139). This approach is not sufficient, although it may form part of a definition of free will. It does not take account of times in which our decisions may be controlled by means other than determinism – for example by drugs or a compulsive mental disorder. In cases such as these, we may be said to not be acting freely.

An alternative compatibilist definition of freedom which does take account of compulsion is offered by A. J. Ayer, who suggests three conditions that must be satisfied for a person to be free: first, if he had chosen to act otherwise he would have acted otherwise; second, his action was voluntary; and third, he was not constrained (Ayer (1954), 282). Ayer's conditions are based on the idea that as long as our desires are an active part of the chain of causation that leads to our actions, then we are free, even though thanks to determinism we do not have any choice about our desires.

I think that Ayer's definition of freedom is a good one, but is not (as he thinks) compatible with determinism. His third condition for freedom, the absence of constraint, is not compatible with determinism, since if determinism is correct we are constrained by the conjunction of the past and natural laws. Ayer argues that there is an essential difference between our acts being compelled and our acts being causally determined. He locates this essential difference in the fact that other kinds of compulsion force us to go against our better judgement (sometimes called our higher order desires), whereas determinism does not entail this. Ayer is correct here, but this distinction does not do the work he needs. Determinism does not force us to go against what we want simply because, if it is correct, what we want is also determined. If anything, I think that this is a greater compulsion than the other kind. Ayer's definition of freedom may be an alternative to the 'could have done otherwise' definition, but it does not seem to me to be compatible with determinism.

At this stage, the compatibilist may protest that it is a special *kind* of cause which interferes with our freedom. I am mistakenly conflating causation and compulsion. For example, it may be that only causes which have no effect on our desires count as constraining. Suppose I want you to remain still. If I were to chain you up, I have bypassed your desires and forced you to remain still. If instead I were successfully to persuade you to stay still, the cause of your staying still involves your desires. In both cases I cause your staying still. However, in the second case your desires are involved and so you are not constrained.

I think that there are problems with this approach. The example above seems to me to confuse freedom of will and freedom of action. It is usually the case that free will and free action go together, but in the case of a physical constraint such as chains, they are separated. In the first case your freedom of action is restricted – but not your will. To see this, consider that you can have entirely different responses to being chained up. You may struggle helplessly against the chains, or you may in fact be perfectly happy to sit where you are – you never wanted to move in the first place. Your will to move (or not) is still free. What is restricted is your ability to effect your will. If it was the case that the chains restricted your free will because the restriction does not 'pass through' your desires, then so too would many facts about the world which we do not normally see as affecting our will (Albritton (2003), 410–412). For example, the fact that gravity prevents me from jumping 20 feet into the air would count as just as much of a constraint upon my will as the chains.

The other problem for the compatibilist comes when it is asked why we should not consider all causes constraining. Why should it only be a special kind of cause, and not all causes, which constrain? The typical answer has its roots in Hume's theory of causation as constant conjunction (Hume (2000), bk. 1, pt. 3, sec. 14). The compatibilist says that causation is simply a regular series of events. To say that A causes B is to say that B has always been observed to follow A; it is not to say

https://doi.org/10.1017/S0034412513000401 Published online by Cambridge University Press

that A exerts some kind of force over B such that B occurs. The idea that causes are constraining relies on this mistaken view that a cause has power over its effect. Here is Ayer on the relationship between cause and effect: 'there is an invariable concomitance between the two classes of events; but there is no compulsion, in any but a metaphorical sense' (Ayer (1954), 282). I think that this answer leads the compatibilist into difficult ground already occupied by the incompatibilist. It seems to destroy or at best weaken the link between the agent and the action (Russell (1988), 316–318). If a cause does not have power over its effect, then it seems that those actions which are not compelled are not connected to the agent – at least, not in the sense which we would usually think important for an action to 'belong' to someone. The compatibilist says that constrained actions are those without a causal connection to the agent's willing, whereas non-constrained acts do have such a connection. Unfortunately, when it comes to explaining why constraint only occurs when a certain type of cause is involved, the compatibilist responds with an explanation of causation which makes it difficult to claim that there can be any real connection between an agent's willing and their actions.

I think that the proposed separation between causation and constraint creates more problems for the compatibilist than it solves. It is worth noting that even if it does succeed the theist may find the compatibilist position difficult to accept. If free will is compatible with our actions being fully determined, then God could stop us from sinning without thereby impinging upon our free will. This would seem to render the free-will defence useless, since God could prevent evil while preserving the good of free will. In any case, I think that van Inwagen is correct, and that compatibilism is mistaken. I now turn to the second stage of his argument for the inexplicability of free will – an attack on incompatibilism.

Van Inwagen's initial criticism of incompatibilism is similar to my earlier criticism of a possible non-causal account of the origin of evil. Incompatibilism says that in order to be free, an agent's actions must be undetermined. But if they are undetermined, it seems that how the agent acts is a matter of chance – and if her actions are matters of chance, they do not seem to be free. One form of incompatibilism which seeks to avoid this criticism is called agent causation. Agent causation is committed to the idea that substances as well as events can be causes. It holds that in acting, the agent is the cause both of their action and also of their willing to perform that action (Clarke (2008), sec. 3.1, para. 2). By identifying the agent as the cause of his or her own willing, agent causation seeks to avoid the criticism that incompatibilism makes our actions products of random chance without committing itself to determinism. Van Inwagen uses an example to show that although agent causation may be correct, it does not show that our actions are not products of chance.

His argument runs as follows: suppose that at a particular time Alice has a choice between lying and telling the truth, and roughly equal inclinations

towards both. She then proceeds to be the agent cause of her telling the truth. Now imagine that God 'rewinds' time back to before Alice told the truth. The conditions are the same, but since she is the sole cause of what she does she is able to make a different decision. This time, she is the agent cause of her telling a lie. Suppose that God rewinds and replays the event 1,000 times, and Alice lies or is truthful in roughly equal amounts. Now if we were asked to predict what Alice will do after the next rewind, van Inwagen says that we will be inclined to conclude that it is simply a matter of chance. Nothing that we or even God could know can tell us anything about what the next choice will be. Even though Alice is the agent cause of what she does and what she wills, it does not seem that this prevents her willing being a matter of chance – in which case, she is not properly free.

Megan Griffith argues that van Inwagen's argument relies on being unclear about the meaning of 'chanciness' (Griffith (2005), 266). The reason that Alice's will and actions appear to be a matter of chance is that they are unpredictable. No one, not even Alice, can know beforehand what she will do. In this sense her actions are indeed chancy. Griffith points out that the reason chance causes a problem for the incompatibilist is that it seems to indicate a lack of control. If agent causation is right, van Inwagen's example shows that our actions are unpredictable, but it does not show that they are uncontrolled (and hence not free). I think that Griffith is right here, but her criticism does not do the work she needs. Van Inwagen does not need to prove that agent causation must be incorrect (although he may think that he has in fact done so). All he needs to do to support his conclusion that free will is a mystery is to show that we cannot know whether or not agent causation is correct. Griffith is right that Alice may in fact be in control of her will and actions in every example. However, what Van Inwagen has shown is that uncontrolled actions by Alice will look exactly the same as controlled, agent-caused acts by Alice and we have no way of telling the difference. It is just as plausible to conclude that Alice's actions are also chancy in the sense of being uncontrolled, or that they are free in some other, unexplained way. There is no evidence to support the conclusion that her actions and will are agent-controlled rather than a product of random chance. I (and van Inwagen) can acknowledge that they may be controlled by agent causation – but since there is no way to know, I think that it is fair to say that free will is still a mystery.

I think that I have shown it is possible to mount a plausible defence of van Inwagen's free-will mysterianism. This position provides the required 'radical inexplicability' for the NEA that I mentioned earlier. It does not offer a cause of free will, but neither does it deny the possibility of one. If free-will mysterianism is adopted, Augustine is able to say that the first evil will is fundamentally inexplicable. This is not because evil is inexplicable, but because all acts of willing (evil or not) are inexplicable – that is, it is impossible to determine whether they are random or agent-caused. The fall is an act of willing and therefore

it is inexplicable. This revised form of the NEA is successful in denying God's culpability for the origin of evil.

I began by describing the problem of evil and Augustine's response to it. I then turned to the further problem that it seems impossible to explain the origin of evil without implicating God in some way. I looked at Augustine's different attempts to answer this criticism and showed that the NEA is the most plausible. For the NEA to work, an explanation for the inexplicability is required. I showed that attempts to locate the inexplicability in the nature of evil fail, and proposed locating it in the nature of free will instead. This is made possible by van Inwagen's argument that free will must be a mystery, since all possible explanations of free will have serious flaws. Finally, I defended van Inwagen's argument against both compatibilist and incompatibilist criticisms. With some modification, Augustine's no-explanation answer to the problem of the origin of evil is effective.

References

Albritton, Rogers (2003) 'Freedom of will and freedom of action', in G. Watson (ed.) *Free Will* (Oxford: Oxford University Press), 408–423.

Augustine (1955) *De Libero Arbitrio*, D. M. Pontifex (ed. & tr.), Ancient Christian Writers: The Works of the Fathers in Translation 22 (New York: Newman).

(2003) *Concerning the City of God Against the Pagans*, H. Bettenson (tr.) (London: Penguin).

Ayer, A. J. (1954) 'Freedom and necessity', in A. J. Ayer (ed.) *Philosophical Essays* (London: Macmillan), 271–284.

Brown, Robert F. (1978) 'The first evil will must be incomprehensible', *Journal of the American Academy of Religion*, **46**, 315–329.

Chappell, T. D. J. (1994) 'Explaining the inexplicable: Augustine on the fall', *Journal of the American Academy of Religion*, **62**, 869–884.

Clarke, R. (2008) 'Incompatibilism', in E. N. Zalta (ed.) *The Stanford Encyclopaedia of Philosophy*, http://www.plato.stanford.edu/entries/incompatibilism-theories/#3.1 (12 January 2011).

Dennett, Daniel (1984) *Elbow Room* (Oxford: Clarendon Press).

Dowe, Phil (2001) 'A counterfactual theory of prevention and "causation" by omission', *Australasian Journal of Philosophy*, **79**, 216–226.

Duff, R. A. (1998) 'Responsibility', in E. Craig (ed.) *Routledge Encyclopedia of Philosophy Online*, http://www.rep.routledge.com/article/L085SECT2 (7 February 2013).

Evans, G. R. (1982) *Augustine on Evil* (Cambridge: Cambridge University Press).

(1999) 'Evil', in A. D. Fitzgerald (ed.) *Augustine through the Ages: An Encyclopedia* (Cambridge MA: Eerdmans), 340–344.

Frankfurt, Harry G. (1978) 'The problem of action', *American Philosophical Quarterly*, **15**, 157–162.

Griffith, Meghan E. (2005) 'Does free will remain a mystery? A response to van Inwagen', *Philosophical Studies*, **124**, 261–269.

Hansson, S. O. (2003) 'Ethical criteria of risk acceptance', *Erkenntnis*, **59**, 291–309.

Hume, David (2000) *A Treatise of Human Nature* (Oxford: Oxford University Press).

Mackie, J. L. (1990) 'Evil and omnipotence', in M. M. Adams & R. M. Adams (eds) *The Problem of Evil* (Oxford: Oxford University Press), 25–37.

Mann, William E. (2001) 'Augustine on evil and original sin', in N. Kretzmann & E. Stump (eds) *The Cambridge Companion to Augustine* (Cambridge: Cambridge University Press), 40–48.

Matthews, G. B. (2004) 'Augustine', in E. Craig (ed.) *Routledge Encyclopedia of Philosophy Online*, http://www.rep.routledge.com/article/B009SECT9 (6 January 2011).

Murray, M. (2005) 'Leibniz on the problem of evil', in E. N. Zalta (ed.) *The Stanford Encyclopaedia of Philosophy*, http://www.plato.stanford.edu/entries/leibniz-evil/ (10 January 2011).

Pereboom, Derk (2004) 'The problem of evil', in W. E. Mann (ed.) *The Blackwell Guide to the Philosophy of Religion* (Oxford: Blackwell), 148–170.
Russell, Paul (1988) 'Causation, compulsion and compatibilism', *American Philosophical Quarterly*, **25**, 313–321.
Tilley, Terrence W. (2000) *The Evils of Theodicy* (Eugene OR: Wipf & Stock Publishers).
van Inwagen, Peter (1983) *An Essay on Free Will* (Oxford: Oxford University Press).
(2002) 'Free will remains a mystery', in R. Kane (ed.) *The Oxford Handbook of Free Will* (Oxford: Oxford University Press), 158–177.