



# RATIONAL *Realm*

*Exploring a rational approach to knowledge and life*

## Free Will and Compatibilism

Leslie Allan

Published online: 11 June 2016

Copyright © 2016 Leslie Allan

The author mounts a case against the libertarian and hard determinist's thesis that free will is impossible in a deterministic world. He charges incompatibilists with misconstruing ordinary 'free will' talk by overlaying common language with their own metaphysical presuppositions. Through a review of ordinary discourse and recent developments in jurisprudence and the sciences, he draws together the four key factors required for an act to be free. He then puts his *4C theory* to work in giving a credible account of how caused human agents 'could have done otherwise'. The author concludes by showing how the deterministic compatibilist view actually underpins our notion of moral responsibility.

To cite this essay:

Allan, Leslie 2016. Free Will and Compatibilism, URL =  
<<http://www.RationalRealm.com/philosophy/metaphysics/freewill-compatibilism.html>>

To link to this essay:

[www.RationalRealm.com/philosophy/metaphysics/freewill-compatibilism.html](http://www.RationalRealm.com/philosophy/metaphysics/freewill-compatibilism.html)

Follow this and additional essays at: [www.RationalRealm.com](http://www.RationalRealm.com)

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sublicensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms and conditions of access and use can be found at [www.RationalRealm.com/policies/tos.html](http://www.RationalRealm.com/policies/tos.html)

## 1. Introduction

Given the modern scientific view of the world and of our place in it, it seems as if we are part of the natural universe. It appears that our hopes, desires and behaviours are just as much subject to the natural laws of the universe as ping pong balls and planets. Consequently, our concepts of moral and legal responsibility and our freedom to choose have been subjected to much critical reappraisal during the last 100 years. Hard determinists have concluded that we should give up on our concept of free will and, perhaps also, on our ideas of moral responsibility. Libertarians, on the other hand, have pushed back hard against the scientific view of mind and human behaviour and maintained that we should preserve our concept of free will while abandoning the mechanistic view of human beings.

What I want to do in this essay is to sketch out a view of human freedom in which we both keep our notion that we sometimes act freely and the view that we are highly complex deterministic beings. On this compatibilist view, free will and determinism are seen to be mutually complementary and not contradictory, as incompatibilists contend. I'll begin by outlining what I think common folk mean when they say that an agent exercised their free will. Next, I'll trace briefly how the meaning of free will has adapted to improvements in our scientific understanding. Then, I'll consider how statements about free will should be understood in terms of counterfactual conditionals that elucidate our feeling that the free-willing person could have done otherwise.

I recognize that in some domains of enquiry, the term 'free will' is explicitly used in the sense of a will that is not completely determined by prior physical states and events. Theologians and libertarian philosophers use the term 'free will' in this contra-causal sense of a will with no sufficient physical cause. This is a technical use of the term in use within these specialist disciplines. What I'm interested in, though, is the meaning attributed to the term in common parlance. I grant that some ordinary language users subscribing to a particular thought-out philosophical outlook (such as theism and existentialism) use the term in this contra-causal sense. What I propose to show is that this sense does not underpin the use of the term in common discourse. During the course of the discussion, it should become clear that incompatibilists are injecting the ordinary language meaning of 'free will' with their own metaphysical presuppositions.

For the purposes of this essay, I will use terms in the following senses. By 'determinism', I mean the view that for the universe, at the macro-level of microbes and humans, every event, including human behaviours, has a set of sufficient physical causes. What I mean by this is that a complete description of the initial state of a closed macro-level system, conjoined with the laws of physics, logically entails future states of the system.<sup>1</sup> By 'indeterminism', I mean the contrary thesis; that there are some macro-level events in the universe that do not have a sufficient physical cause.

---

<sup>1</sup>I include within the determinist's thesis the possibility of indeterminacy at the quantum level as indeterminacy at this level does not impact causal inferences at the macro-level. Quantum physicists tell us that the quantum wave function collapses as soon as a quantum particle interacts with another particle, such as in any biological system.

By 'compatibilism' (or 'soft determinism'), I mean the thesis that determinism is true and is compatible with the view that human beings possess 'free will', as ordinarily understood. By 'incompatibilism', I mean the view that the thesis of determinism is not logically compatible with the notion that human beings possess free will. Incompatibilists can either accept the thesis of determinism while rejecting the notion of free will (in which case they are 'hard determinists') or reject the thesis of determinism while accepting the view that we possess free will (in which case they are 'libertarians'). Finally, in this essay, when I write about a 'free act', I mean an intentional act that is the result of the exercise of a free will.

## 2. What Is Free Will?

In this essay, I am interested in exploring whether the notion of ‘free will’ is compatible with the belief that all of our actions are determined by a network of physical causes that extend back in time to a time even before that of our own birth. This is the deterministic picture painted by modern science, from the fields of neurology, information theory, evolutionary psychology, genetics and a number of other connected fields. Given that we are highly complex physical structures obeying the laws of physics, can we still legitimately describe some of our most important choices and actions as being manifestations of our free will? Much of the answer to this question hangs on what we mean by a person exercising their ‘free will’.

Regrettably, much of the philosophical debate on what ‘free will’ means, as well as the debate about related terms, such as ‘agency’, ‘control’ and ‘choice’, is divorced from an analysis of how ordinary people use these words in their day-to-day living. I think such an ordinary language analysis is the starting point for an examination of the question of whether free will is compatible with determinism. Incompatibilists argue that determinists ought to stop labelling some actions as being the result of free will. They argue that determinists should modify their day-to-day discourse when it comes to ascribing free agency to some voluntary behaviours, such as choosing which tie to wear to work. To settle this question, enquiring into what common folk take ‘free will’ to mean becomes critical.

I’ll start by reviewing how the terms ‘free’ and ‘free will’ began use within English speaking communities. The term ‘free’ arose from the Old English word ‘freo’ in the thirteenth century. This word meant:

free, exempt from, not in bondage<sup>2</sup>

Between the years 1525 and 1535, the conjoined term ‘free will’ arose for the first time.<sup>3</sup> In the literature of the day and in the ensuing decades, the term was used to denote a person’s will that was not constrained or forced.<sup>4</sup> This meaning of an unencumbered and uncoerced will carries through to the modern era. In addition to the notion that a free will is an uncoerced will, modern advances in science and jurisprudence have led to the recognition that in other situations, a person’s will can also be encumbered or restricted. I’ve grouped the kinds of situations in which a person’s capacity to exercise their free will is restricted into four classes. These are:

1. coercion
2. manipulation
3. addiction
4. mental illness

<sup>2</sup>See, for example, entries under ‘free’ in *Collins English Dictionary* [2012], URL = <<http://www.dictionary.com/browse/free-will>> and *Online Etymology Dictionary* [2016], URL = <<http://www.etymonline.com/index.php?term=free>>.

<sup>3</sup>See the entry under ‘free will’ in *Dictionary.com Unabridged* [2016]. Source location: Random House, Inc., URL = <<http://www.dictionary.com/browse/freewill>>.

<sup>4</sup>Examples of common use are given in the entry under ‘free will’ in *Wiktionary, The Free Dictionary* [2016], URL = <[https://en.wiktionary.org/w/index.php?title=free\\_will&oldid=36319568](https://en.wiktionary.org/w/index.php?title=free_will&oldid=36319568)>.

I will now discuss each of these classes and, along the way, illustrate each with representative examples of how they appear in common discourse about freely willed behaviours.

## 2.1 Coercion

The various forms of coercion are the earliest recognized encumbrances on a person's ability to act freely. For example, in Shakespeare's *Antony and Cleopatra* [c. 1606], Octavia pleads to Caesar that he travelled to Rome of his own free will in spite of the constraints put on him by Mark Antony. As Octavia puts it, 'Good my Lord, To come thus was I not constrain'd, but did it On my free-will' [3.6.65].

In Charles Dickens' *Nicholas Nickleby*, Nicholas begs the young Madelaine to not marry Mr. Gride as she is being unknowingly impelled to do so. Madelaine protests, 'I am impelled to this course by no one, but follow it of my own free-will. You see I am not constrained or forced by menace and intimidation' [1843: 313f].

From these early uses, note in particular how a free will is contrasted with a constrained or forced will. It is not contrasted with an uncaused will. The constraints identified by these characters are constraints put up by other people and not causes lurking in heritable characteristics or brain physiology, as some incompatibilists argue.

To Dickens' coercive means of menace and intimidation, we can add threats of violence, hostage taking, threats to reputation and withdrawal of critical resources. Robbery at gunpoint and forced marriages are paradigmatic examples of the prevention of the exercise of free will. I will say more on paradigmatic examples in §3 below.

This idea of absence of coercion as being central to the notion of 'free will' carries through to the modern day. The *Collins English Dictionary* renders one of the two meanings of 'free will' as:

the ability to make a choice without coercion: *he left of his own free will: I did not influence him*<sup>5</sup>

Similarly, the *Webster's Revised Unabridged Dictionary* gives the first of two meanings to 'free will' as:

A will free from improper coercion or restraint. "To come thus was I not constrained, but did On my free will."<sup>6</sup>

Following are some modern day examples of ordinary people identifying the absence of free choice with coercion. In the United Arab Emirates, many women are trafficked into prostitution. The Director of Ewa'a shelters for trafficked women and children said to a news reporter, 'How many women do this work of their own free will and how many are coerced is unknown' [Lageman 2016].

<sup>5</sup>See the entry under 'free will' in *Collins English Dictionary – Complete & Unabridged 10th Edition*, HarperCollins Publishers, URL = <<http://www.dictionary.com/browse/free-will>>.

<sup>6</sup>See the entry under 'free will' in *Webster's Revised Unabridged Dictionary*, G & C. Merriam Co., 1913, ed. Noah Porter.

In a newspaper article on marriage, authors Emmanuel and Moussa [2015] explain the difference between an arranged marriage and a forced marriage.

In an arranged marriage, the meeting of the spouses is arranged by family members, relatives or friends, but the spouses agree by their own free will to marry. In contrast, in a forced marriage, one or both spouses are coerced into the marriage – the union takes place without their freely given consent – either under duress, threats or psychological pressure.

Consider a final case of armed kidnapping in the United Kingdom. Colombian drug criminals forced two young women at gunpoint into smuggling cocaine through customs. After the women's arrest, the Archbishop of Lima claimed they 'had no choice but to follow orders'. He added, 'If they have been coerced or threatened as I think they are going to argue, then the fact that they physically had it in their possession may not mean that they were intentionally or wilfully doing it.' The father of one of the women pleaded that his daughter 'would never do anything like this of her own free will' [Sky News 2013].

## 2.2 Manipulation

Whereas coercion gives rise to a feeling of psychological pressure in the agent, manipulation does not. Manipulation includes more direct means of mind control, such as hypnosis, brainwashing, brain implants and zombie drugs. As these items and practices surfaced only in the modern era, the original notion of an unfree will as a coerced will needed to be extended. I will say more on this in §5 below. Here, I will give some illustrative examples of how non-philosophers and non-theologians regard free will in these cases. The basic notion remains the same: an unfree will is an encumbered will.

Father Edwin Healy countenanced against the practice of hypnotism because it ‘deprives the subject of the full use of reason and free will’ [cited by Catholic Education Daily 2015].

In a piece on mass hypnosis in *Before It’s News* [2012], the author explains:

You think that you’re in control of your life, that nothing is capable of influencing you without you first judging it, little do you know that you are being hypnotised by many different organisations on a daily basis. The actions you think are of your own free will, are actually influenced by these external organisations.

Whether people under hypnosis are really deprived of their free will or not is not the issue here. As these two examples demonstrate, for ordinary folk who think it does, for them the absence of free will under hypnosis is tied to the person’s lack of reasoning and judgement in particular. No deliberation is given to whether the person under hypnosis has uncaused brain states, as the incompatibilist maintains.

Judgements made about brainwashing are along similar lines. Yeonmi Park fled Kim Jong-il’s North Korea with her parents. She claimed she was brainwashed by the regime with the result that ‘I had not been a real person – I was created for the regime to work for them. If they ordered us to die, I would’ve died for them. I wasn’t a human – I was something else.’ After escaping and educating herself, she said, ‘I now have free will’ [SBS 2014].

In the case of the Commonwealth of Virginia v. Lee Boyd Malvo, 17 year old Lee Malvo was charged as an accomplice in a string of random sniper killings. An expert witness at the trial, Dr Steve Eichel, said of Malvo afterwards that he ‘was indoctrinated (“brainwashed”) into his role as John Muhammad’s loyal co-perpetrator [and] that Lee was not capable of freely forming an intent . . .’ He added, ‘Lee’s “old” self (a highly vulnerable boy who was and is quite bright, personable and troubled by a traumatic past) became engulfed by his “new” self (he even took a new name, as many cultists do, and became “John Malvo”), a pseudoidentity that was capable of committing horrendous crimes for the “cause” of his leader, John Muhammad’ [Eichel 2004].

As these examples illustrate, brainwashing robs a person of their free will through replacing their personal identity, their character, with another. Central also is the idea that this manipulation is done deliberately by another agent. The notion of contra-causality seems inconsequential.



This theme of third-party control and loss of character appears also with judgements about brain implants. Although mind control through brain implants is still very much the preserve of science fiction, some believe it is happening now. Hodges [2015] writes that the government's aim in microchipping every citizen is 'the complete control of every individual through mind control' and 'to control all thought, all emotion and consequently, all behavior. The end result will be to remove all potential opposition (i.e. free will).' For Hodges, we all currently possess free will even though our minds are strictly subject to causal laws. The tipping factor for this writer is third-party control and not determinism.

This linking of the external control wrought by brain implants to the loss of free will is also reflected, for example, in the writings of Johnston [2002], West [2013] and Jeffery [2016]. Note that both Hodges and Johnston are committed Christians who think that free will is compatible with a mind enmeshed in physical causes. Johnston's [2002] idea that free will is 'the ability that humans possess that allows them to make decisions based upon their own deductive and reasoning skills' is an idea that I will return to in §4 below.

The robbing of free will in the case of brain implants, however, is not so clear cut. As Glannon [2014] has pointed out, brain implants using deep brain stimulation (DBS) can restore the ability to choose freely in some patients suffering certain debilitating diseases. These include Obsessive-compulsive disorder (OCD), Alzheimer's disease and major depression. In these cases, the brain implant can bring back the patient's old self that was lost to the disease. The restoration of free will for these patients hinges on the concept of free will illustrated in the previous types of encumbrances; that is, the reinstatement of the patient's character and ability to reason over their actions.

## 2.3 Addiction

Various forms of addiction are regarded as psychological compulsions that inhibit the exercise of free will. These include alcohol, substance, work and gambling addictions. Some addictions are a symptom of a mental illness, which I deal with in the next section. Examples of these are addiction to sex, hoarding, kleptomania and pyromania. For the ordinary person on the street, as well as for medical and legal experts, many addictions are seen as compromising a person's ability to choose freely. Here are some examples of these opinions.

When asked about the drug addict's lack of choice and what this means for free will, the director of the National Institute on Drug Abuse, Dr. Alan Leshner, explained:

Most people are able to control their initial drug use. They're able to exert their will over it, but once they are addicted, it's a myth that many people just decide to break their addiction. . . . They need help to deal with the compulsive, uncontrollable drug use.

[Moyers 1998]

Leshner continues:

You are an addict because your brain has been changed by drugs. You're in a state where the drug has totally taken over your being. . . . So, there's something about these biological changes that are going on at the cellular level that gets translated into compulsive, uncontrollable drug use on the behavioral level.

[Moyers 1998]

What's important to note is that while Leshner thinks that one's 'brain is constantly changing as a function of the experiences one has', the addicted person's brain undergoes 'a very dramatic change' [Moyers 1998]. Free will, for Leshner, is not about having some of one's brain states form independently of one's genetic constitution and environment. It's about being free of psychological compulsion and being true to one's character; to one's being.

Dr. Denise Cummins mirrors Leshner's view that drug addiction robs the addict of the 'free will to choose' not by subjecting previously uncaused brain states to physical causes, but by changing the causes that activate the reward circuitry in the addict's brain. For Cummins [2014], the loss of free will from drug addiction is from the drugs 'power to enslave us'.

According to the Law Library, a judge will only allow an accused person to enter a guilty plea if they consider that the accused 'exercised free will'. The guilty plea is not accepted if the 'defendant isn't mentally competent at the time he agrees to the plea, for example, due to a developmental disability, intoxication or influence of narcotics'. Once again, the ability to reason about one's actions figures prominently in the consideration of whether the act is free. The accused must be able to understand the court proceedings and 'consult with his lawyer with a reasonable degree of rational understanding' [Law Library [2016].

Here, the point is not whether an addiction really limits the exercise of free will. Opinions will vary considerably. What these examples demonstrate is that the common person and the professional alike link the loss of free will in cases of addiction to feelings of compulsion, loss of personal identity and reasoning ability. The incompatibilists' notion of contra-causality is notably absent from these considerations.

## 2.4 Mental Illness

A mental incapacity to reason and regulate one's behaviour, either because of genetic history, accident or disease, is regarded as an important restrictor of an agent's free will. Here, I give some examples from the generalist literature.

Following the tragic incident in which psychiatric patient, Gulchekhra Bobokulova, beheaded the young child she was minding for her parents, Ms Bobokulova's father was reported as saying that she would never have committed such a crime of her own free will. He added, 'She needs treatment. She'd never have done this herself' [Oliphant 2016].

June experiences hallucinations that command her to commit various acts. She writes, 'I find that it is like my free will has been removed and have no alternative but, to obey.' She also links this feeling of compulsion to the feeling that the voice inside her head 'keeps me from being myself' [Nemade 2009].

Herschel Hardin is the father of a child with schizophrenia. He laments that the illness deprives all those affected by the illness of the capacity for free will by robbing them of rationality and autonomy. 'Their personalities are subsumed by their distorted thoughts' [Hardin 1993], as he puts it.

As these examples show once again, the common themes in the ordinary person's ways of thinking about free will are that its presence requires that a person's character is intact and that they can reason. What is also of critical significance here is the fact that the insanity defense has a long history in jurisprudence. In many legal jurisdictions, the defense applies when it is judged that the accused is dispossessed of their free will. Experts for the defense must testify that the accused is either cognitively incompetent, unable to comprehend the nature of the act and to reason about it, or volitionally incompetent, unable to control their impulses. In either case, expert witness for the prosecution is not called upon to demonstrate what the libertarian asserts; that the causes of the defendant's transgression were themselves uncaused activities in the brain. One such example of how the absence of free will in the mentally incapacitated works in a legal defense is *Colorado v. Connelly* [1986]. For a useful history and critical review of the insanity defense, see Grachek [2006].

### 3. Paradigmatic Examples of 'Free Will'

Although the examples of restrictions on free will given in the previous section are not exhaustive, they do represent the way ordinary folk think and talk about free will. Readers can gain more insight into folk thinking by reading comments from internet forum contributors on the freedom-restricting circumstances mentioned. The key point I am making here is that a survey of generalist literature and discussion forums reveals a paucity of thinking about the loss of free will as a loss resulting from certain brain states abandoning their acausal status and becoming fully engaged with other causal physical processes in the brain. What limits free will in ordinary people's minds is not causality, but coercion, manipulation, addiction and mental illness. This way of thinking has only been further elaborated and refined by modern developments in psychiatric practice and jurisprudence.

My point can also be illustrated by examining paradigmatic uses of the term 'free will' in common discourse. Consider this dialogue about a recent marriage.

Fred asks John: 'Did you marry Kate of your own free will?'

John answers: 'No. If I did not marry Kate, my entire family would have abandoned me and I would have been evicted and left with no money.'

Alternatively, John could have answered: 'Yes. We fell in love at university and my parents had no objections.'

In either case, Fred's query is satisfied and he leaves the conversation knowing whether John married freely or not. Crucially, Fred's question and John's answer hinged on whether the marriage was coerced or not by John's family.

Now consider a second dialogue about handing over a wallet.

Mary asks Peter: 'Did you give your wallet to that man of your own free will?'

Peter answers: 'No. He was holding a gun to my head, threatening my life if I did not do so.'

Alternatively, Peter could have answered: 'Yes. I wanted to go for a swim in my shorts and did not want my wallet to get wet.'

Again, Mary's question is answered with reference to whether the act was coerced or not. For Fred and Mary, their question was not a neurophysiological question about the causes and causal absences happening in John and Peter's brains. Their question was a practical question, rooted in their day-to-day lives.

Hard determinists have taken our modern, scientific understanding of the brain and overlaid this causal model onto what they think common language terms, such as 'free will', mean. Scientists, such as Sam Harris, have also fallen into this trap of injecting their metaphysical understanding of the world into what they think is the common person's use of ordinary language terms. This error is akin to scientists, on discovering that atoms are mostly empty space, proclaiming that tables are not 'really' solid. A similar mistake occurred with our notion of a 'coloured object'. Physicists tell us that atoms are not coloured. It turns

out our perception of colour is a function of how our brains process light of various electromagnetic wavelengths. On the basis of this insight, some philosophically naïve scientists seek to correct our common sense notion by proclaiming that apples are not 'really' red. I see hard determinists making a similar mistaken attempt to overlay our scientific understanding of human beings onto ordinary, day-to-day discourse.

An incompatibilist may object that the examples I give showing how coercion, and the other various circumstances I bring up, limit free will only serve to reveal what people think are the behavioural and environmental correlates to deterministic brain states. This goes no way to disproving, the objection continues, that common people regard a free agent as one whose brain states are not completely determined by prior physical causes.

I think a cursory review of generalist literature and listening to 'free will' talk shows this objection to miss the mark. Take, for example, the debate conducted on Debate.org [2016] on whether an addict still has free will while addicted. Those saying yes, they do, appeal exclusively to the fact that addicts can and do make choices.

Here are two more examples. Gambling addict, Chris Wright, argued that addicts like him retain the capacity of free will, enough to regain a sense of responsibility and a 'degree of agency' to choose otherwise [Wright 2013]. He makes no mention of uncaused causes.

Journalist Andrew Brown [Brown 2016] argued that obesity is not an addiction that takes away a person's free will as addicts change their behaviour in response to improved social circumstances. For Brown, freedom of the will is about choice and not brain chemistry. In view of these examples, I think it incumbent on incompatibilists who believe that ordinary folk think in terms of contra-causality to demonstrate their thesis with common language examples.

The 'common language' critique I advance here can be extended to judicial language and thinking. Throughout the modern history of jurisprudence, in determining whether a defendant was absent of the capacity of free will at the time of the crime, no jury or judge has requested or called in expert witnesses to attest to the fact that at the time of the crime the defendant's relevant brain states transitioned from a physically contra-causal state to a causal state. This is not surprising as there is no dualist theory evidencing the circumstances in which neuron firings in the brain get removed from the chains of causation to which other neurons belong.

In fact, judges examine, and juries are asked to consider, whether there were any circumstances that either eliminated or mitigated the defendant's ability to choose freely. The types of circumstances that the judge and jury consider include precisely those types of encumbrances outlined above: coercion or manipulation by a third party, drug addiction and mental illness. These are precisely the impediments to free will to which the compatibilist points. For a systematic summary of judicial defenses, see, for example, Robinson [1982].

## 4. Four Necessary Conditions for Free Will

In the previous section, I crystallized the four types of situational impediments to the exercise of free will: coercion, manipulation, addiction and mental illness. What is it about these situations that minimize a person's capacity to act freely? In my treatment of these situations, four requirements for an act to be called 'free' recurred throughout. For brevity, let's call this compatibilist account of the requirements for free will the '4C theory'. These 4Cs are:

- (1) Compulsion
- (2) Control
- (3) Character
- (4) Cognition

(1) The first requirement is that the act not feel compelled by the agent's situation. The feeling of compulsion I am referring to here is an introspective psychological experience. Here, the agent feels that they will sacrifice something of great value to them if they do not act in a particular way. The agent feels that they had no choice but to act as they did. Note that intentional acts that are coerced are but a subset of compelled acts. Some compelled acts are not coerced in the sense that the agent is not forced to comply by a threatening third party, but is nonetheless felt compelled to act as they did.

(2) The second requirement is that the act not be controlled by a third party. Unlike in the case of compulsion, here, even though the agent's thoughts and actions are being controlled, they do not feel as if they are being compelled by circumstance. However, with their actions being manipulated either directly or indirectly by a third party, they have lost their autonomy. This requirement goes to the heart of what it is to be a moral agent with responsibility for one's actions. When control of a person's behaviour is surrendered to another moral agent, the locus of responsibility moves along that line of control to the third-party agent in control of the human puppet's behaviour.

(3) The third requirement is that the action is consonant with and a consequence of the agent's character. When the agent's behaviour is out of character, the person is not a *bona fide* agent of their own actions. This requirement often acts in tandem with (2) the requirement for lack of third-party control as a marker of personhood.

(4) The fourth requirement is that the agent has the cognitive capability to offer reasons for their action and to deliberate about alternative courses of action. Without rational agency, the person is not exercising autonomy and is better described as a passive repository of impulses.

Each of these four requirements must necessarily be satisfied for a particular act to be considered the act of a free agent. A good test of the necessity of each of these requirements is to measure them against actual moral and legal deliberations, such as the ones illustrated throughout this essay.

What ties all four requirements together is the fundamental axiom expressed in §2; that a free will is an unencumbered will. With the advent of scientific knowledge and modern technology, this basic understanding of encumbrance as compulsion has been supplemented with these additional requirements for moral and rational autonomy. Various compatibilists have picked up one or more of these central requirements and developed complex theories around them.<sup>7</sup> My view is that a comprehensive and compelling account of what it is to exercise free will must do justice to all four requirements. I will expand on this idea in the next section.

What seems clear is that philosophical and legal thought over the last century or so has largely coalesced around the view that freedom of the will is a characteristic of an autonomous, conscious agent who can reason and deliberate about alternative courses of action. The thinking here is that such a person is constituted by their character and that within the bounds of this character, the agent faces a range of options on how to act in a given situation. When this range is encumbered or restricted by either subverting the person's character or compromising their capacity for rational deliberation and action, the person's freedom is diminished.

---

<sup>7</sup> See McKenna *et al* [2015] for a comprehensive survey of compatibilist approaches.



## 5. Evolution of Free Will Concept

As a lead-in to a discussion about how the concept of free will has evolved in the face of societal changes, I want to clarify some other aspects of 'free will' talk that can lead to misconceptions. The first is that when we speak of a person having free will, its possession is not an all or nothing affair. *Contra* extreme libertarians, such as Sartre and Descartes, some intentional acts are free and some are not. Handing over one's wallet when a thug holds a gun to one's head is an intentional act, but it is not freely chosen. Choosing tea or coffee at a friend's house, on the other hand, is.

Allied with this, some intentional acts in particular admit of shades of grey. The act of robbing a store to feed the addict's heroin habit may have been neither completely free nor completely unfree. How free it was depends on the particular nature of the person's addiction at that time and of their environment. As Dr. Alan Leshner points out when speaking of the addict's switch from non-addiction to addiction, 'that switch probably moves gradually, not precipitously' [Moyers 1998].

This idea of the non-exclusive and graduated nature of free will helps explain why reasonable people can and do disagree about which actions are freely willed and to what extent. This is in large part due to the fact that 'free will' originated as a folk psychological concept and remains so in common parlance. It is too much to expect that folk psychology enjoys the same amount of precision as scientific concepts, such as 'mass' and 'time'.

Another spur to disagreement is the evolution of the concept of 'free will' itself. Advances in neuroscience, medicine and jurisprudence over the last century raised new questions about the nature and extent of free will. To what degree do particular brain injuries impair the patient's exercise of free will? What of drug addiction and brainwashing? Questions such as these have exercised medical researchers, philosophers and judges over many decades. The enquiry and debate continues at the cutting edges of technology, with experts disagreeing on many important points.

This evolution in our thinking about free will is akin to modern developments in our Western definition of 'death'. Prior to the modern era, clinicians and lay people had a relatively clear and unambiguous understanding of who was dead and who was not. Absence of respiration and pulse (cardiorespiratory death) were, for centuries, the indicators of a person dying. With the advent of modern neurophysiology and improvements in technological capability, we can now distinguish cardiorespiratory death from a persistent vegetative state (PVS) and from brain-stem death.

Most importantly, doctors can now keep a patient's body alive in all three states with cardiopulmonary resuscitation (CPR) and artificial nutrition. As medical technology advanced, the medical profession's capacity to keep bodies alive way beyond the point of cardiorespiratory death increased dramatically. This blurring between our ordinary conceptions of life and death had crucial implications for medical, legal and moral decisions, especially around the practices of organ transplantation and euthanasia. Just as with our notion of 'free will', in the ensuing decades, philosophers, doctors and lawyers have refined

our definition of 'death' to meet the challenges of technology and our increasingly sophisticated understanding of ethics, medicine and psychology.<sup>8</sup>

The key point I want to drive home here is that during this long and complex debate about how we should refine our definition of 'death', no philosopher, lawyer or medical expert argued that because refinements needed to be made, no one is 'really' dead. There were then and still are paradigmatic examples of people who are dead (J.F.K., Gandhi) and people who are not dead (Obama, you, me), just as there remain to this day paradigmatic examples of free acts (choosing coffee over tea) and unfree acts (handing over one's wallet at gunpoint). Advances in our scientific understanding and technological capability force us to refine our definitions by adding qualifications to previous notions of 'dead' and 'free will'. The addition and clarification of these qualifications over time is neither surprising (given the nature of philosophical discourse) nor arbitrary if it helps us makes sense of our intuitions.

Some current diseases pose important questions for our notion of 'free will'. For example, does a person who experiences a benign change in personality with the onset of a brain tumour still possess free will? Some incompatibilists push these boundaries of our understanding even further with hypothetical scenarios involving created human beings and brain microchips. With these scenarios, incompatibilists expose the difficulties inherent in our notion of 'free will'.

This situation, though, is no different to our difficulties in understanding what is meant by 'dead' when we think of utilizing technologies that surpass what is currently possible. For example, scientists working in the field of cryogenics invite us to think about preserving a person's character and identity after they have experienced death, as usually understood, through a process of delaying what they call the person's 'information-theoretic death'.<sup>9</sup> But once again, just as the semantic, moral and legal questions posed by the possibility of delaying 'information-theoretic death' do not mean that no one is really 'dead', these kinds of questions posed by possible future technologies likewise don't mean that no one really exercises 'free will'.

These questions posed by technology and that require thoughtful answers highlight another important aspect of the debate over meanings. How philosophers, law makers and medical professionals define 'death' in their various communities is influenced by how these thinkers view moral and legal obligations. Their moral frameworks guide, to some extent, where they wish to draw the dividing line between life and death.

The situation is exactly the same with how technological advances impact our ideas about where to draw the line between free and unfree acts. When we think it acceptable to turn off a person's life support and when we think a brain tumour robs a person of their free will depends, to some extent, on our notions of moral agency and what it means to be a person. We can readily admit this complex interplay between semantics and ethics without abandoning entirely our distinctions between 'life' and 'death' on the one hand and 'free will' and 'unfree will' on the other. The upshot here is that this complex interplay is most

---

<sup>8</sup>For a brief overview of changes in the definition of 'death', see *Wikipedia* [2016a].

<sup>9</sup>For more information on the concept and implications of 'information-theoretic death', see, for example, de Wolf [2011] and *Wikipedia* [2016b].

visible at the cutting edges of technology and medicine and do not impact how we regard paradigmatic cases of death and free choice. Because of this mutual dependence between our sense of moral obligation and what hard determinists consider to be the intellectually bankrupt notion of 'free will', some hard determinists canvass the idea that we should jettison ethics altogether. I shall have more to say on this important relationship between free will and moral obligation in §8 below.

## 6. Intuition of Free Will

Some libertarians make much of our intuition that when we act freely, we feel as if we act unrestrained by antecedent physical causes; that we act contra-causally. Campbell [1967: 41], for example, urges us that:

human beings so obstinately persist in believing that there is an indissoluble core of purely self-originated activity which even heredity and environment are powerless to affect . . . because they feel certain of the existence of such activity from their immediate practical experience of themselves

Hard determinists, in contrast, call into question this feeling, labelling it as just an 'illusion' of free will. I think both positions are mistaken. Think about a paradigmatic free act, such as choosing coffee over tea at your friend's house (assuming you enjoy both kinds of beverage). Sure, you don't feel any kind of physical impulse forcing your finger to point to the coffee. You would be the first to know if there were such a force acting on your finger. Furthermore, you don't feel any kind of impulse forcing neurons in your brain to adopt particular action potentials. It's the feeling of absence of such a force that the libertarian cashes out as the feeling of having free will.

But notice the queerness of the latter kind of feeling. Feelings of your finger being forced are natural and we all know what that feels like. However, what would it feel like to have particular neurons in your brain forced into particular states—states that go on to generate forces that determine the states of the motor neurons that ultimately trigger your finger pointing? There are currently billions of neurons in your brain that at this very moment have their states determined by the action potentials of other neurons. And yet you have no feeling of these neurons being forced. If you cannot know the feeling of neurons in your brain being forced into particular states, how can you or anyone else possibly know what it feels like to have some neuronal states underdetermined by other neuronal states? The feeling of free will that we have must be something other than the feeling that some of our brain states are underdetermined by physical forces.

What we can conclude from this introspective psychological reflection about which kinds of states we can have a feeling about and which we can't is that the hard determinists give away too much to the libertarians. The 'illusion' of contra-causal free will that the determinists refer to is an illusion itself. It is a chimera that only serves to confuse the debate. We, in fact, do not have an 'illusion' of contra-causal free will because it's not the kind of thing that we can have an illusion about.

If our feeling of free will is not a feeling of contra-causal willing, then what is it? The answer, I suggest, is the 'feeling' of the absence of compulsion. The feeling of free will, then, is not so much a positive feeling. It is the absence of a feeling. Just as being 'pain-free' is not a feeling. Being 'pain-free' is simply the absence of the feeling of pain.

## 7. Free Will and ‘Could Have Done Otherwise’

### 7.1 Compulsive Acts

In the previous section, I tried to show that, *contra* libertarians, we do not have a feeling of contra-causal free will when we act freely. And *contra* the hard determinists, we also do not have the ‘illusion’ of contra-causal free will. Both of these approaches mistake what it feels like to act freely. I argued that our ‘feeling’ of acting freely is simply the recognition that we are not compelled to act the way we do.

Consider now the other side of the ‘free’/‘unfree’ coin. When we commit an intentional act unfreely, we feel as if we were not able to choose otherwise. In common parlance, we sometimes hear someone complain that ‘I had no choice’ when forced to do something they did not want to do.

One of two police officers threatened with beheading by a lone terrorist, Numan Haider, reported to a coronial inquest afterward that ‘he did not believe the detectives had any other choice but to shoot Haider’ [Dowling 2016]. The Officer added, ‘He was going to slit my throat and cut my head off. I don’t think we could’ve done anything else at all’ [ninemsn 2016]. The Officer’s feeling that he had no other choice and that he could not have done otherwise is not a reference to an unbroken chain of physical causes going on inside his brain, but a feeling of psychological compulsion. His feeling was not one of physical forces acting at the level of his neurons, but one of psychological force in his mind.

As the hard determinists regularly and rightly point out, if a person talking about their unfree act (such as the Officer in the above disturbing scenario) is referring to the existence of sufficient physical causes of their action, then of course the person could not have done otherwise. Clearly, this cannot be to what the Officer in the above scenario is referring. The Officer values extremely highly his own life and the life of his threatened colleague. Given the choice between his and his colleague’s lives and the welfare of a terrorist, his psychological makeup, coupled with the dire circumstances in which he finds himself, compels him to choose to save their own lives.

How then should we understand the truism that the Officer could not have acted otherwise than to shoot the terrorist? The modal past participle, ‘could have’, is a commonly used modal verb in the English language and all other major languages. In this case, a credible ordinary-language compatibilist understanding of this application of the modal verb is to deny the truth of the following conditional:

Given the Officer’s character, the Officer would have done otherwise in this situation if certain external circumstances were different.

This analysis makes use of a common distinction between the intrinsic properties of an entity that give it its capability and the extrinsic conditions necessary for that capability to be actualized. We say, for example, that the electric fan in my study is capable of cooling when it has a working motor, fan blades, power cord, etc. These items constitute its intrinsic properties. We also say that the fan remains capable of cooling even when someone has

switched it off or placed it inside a sealed box. In this case, the fan remains capable of cooling even when one or more of the extrinsic conditions are missing.

Linking this to the case at hand, the intrinsic properties of the Officer is the bundle of characteristics we call his 'character'. By 'character', I mean that combination of beliefs, desires and values that make up a person's psychological profile. I think this folk psychological understanding of 'character' is sufficient for our purposes here, along with all the imprecision that the concept brings with it.

The extrinsic conditions of the situation are those conditions necessary for the Officer to be able to physically shoot the terrorist. The 'external circumstances' of the situation range over features of the situation that are external to the agent's character, but not over the extrinsic conditions that enable the act. For this analysis, it is crucial that this step preserve the opportunity for the Officer to shoot the terrorist. Otherwise, there can be no question about whether his shooting of the terrorist was a free act or not. If his gun was not loaded, for example, the freeness of his act of shooting the terrorist no longer remains an open question. Asking what the Officer would have done in different circumstances, then, is to ask what he would have done under different external circumstances that leave intact his opportunity to shoot the terrorist.

So, given the situation that the Officer faced with his life and the life of his colleague under immediate threat, can we say that there existed factors external to his character that, had they been different, would have led him to do other than to shoot the terrorist? It appears not. If the terrorist had a different appearance or spoke a different language, or if the Officer was working overtime or nearing retirement, it is clear that none of these kinds of differences would have led the Officer to act in a way other than he did. His biologically-based survival instinct, coupled with the defensive skills inculcated through his police training, were decisive elements of his character that overrode any environmental contingencies. We can conclude that the Officer's unfree act in shooting the terrorist can be understood in terms of his not doing otherwise even if the external circumstances had been different. Drawing on the fan analogy above, the Officer's inability to do otherwise is akin to my fan's inability to avoid cooling the room if it is permanently wired to the mains electricity with no off switch. In both cases, they could not do otherwise because of their intrinsic properties.

## 7.2 Analyzing Free Acts

In the previous section, I offered a compatibilist analysis of an unfree act in which the agent felt they could not have done otherwise than what they did. I want to move now to considering the opposite kind of act; that of a free act. The person on the street feels that when they act freely, they could have done otherwise. This common belief forms the basis of a key incompatibilist argument against the compatibilists. The argument states that in a deterministic world, a person could not have done otherwise because the initial conditions (the person's brain states and their physical environment), coupled with deterministic physical laws, logically entail the person's act. The incompatibilist concludes that it was not open for the person to have chosen other than the way they in fact did.

I want to show that this argument is mistaken. Let me start by contrasting two senses of 'could have done otherwise' used in ordinary language. I will call these the *PLIC* and the *PLUC* sense. We use both senses when talking about inanimate objects. The *PLIC* sense refers to Physical Laws + Initial Conditions and is more typically the sense used by scientists when discussing scientific theories and making predictions from those theories. To illustrate this sense, think about this scenario. John is at his local hotel playing billiards with his friends. He hits the billiard ball precisely the right way for it to land in the top pocket. It is natural for us to say, 'The billiard ball could not have done otherwise than to go into the top pocket.' Of course! Because cues, billiard balls and billiard tables obey physical laws. In this case, expanding 'could not have done otherwise' as a *PLIC* counterfactual conditional means:

Given the laws of physics, the characteristics of the billiard ball and the exact same initial conditions, the billiard ball would have landed in the top pocket.

By 'initial conditions' is meant the state of the system under consideration at the start of a process. In this case, the initial conditions include the position, size, weight, velocity, shape, smoothness and hardness of the billiard balls, cue and billiard table at the time of John striking the billiard ball. For this *PLIC* sense, both the characteristics of the billiard ball (that is, its intrinsic properties) and the other initial conditions are made invariant.

Now imagine that on his next shot, John miscues seriously with the result that the billiard ball begins hurtling up into the air in the direction of some hotel patrons standing at the bar. Fortunately, the billiard ball does not hit anyone as no one was standing in the position where the billiard ball came hurtling through. It's natural for us to say, 'The billiard ball could have knocked someone out'. Now, it's the same billiard ball obeying the same physical laws in the same deterministic universe. Yet we are saying that the billiard ball could have behaved differently.

Here, we employ the *PLUC* sense to understand what we mean when we say that the billiard ball 'could have done otherwise'. The *PLUC* sense refers to Physical Laws + Unchanged Character and is more typically employed when discussing events of significant concern to humans. So, in this *PLUC* sense, when we say that the billiard ball could have harmed patrons, we mean:

Given the laws of physics, the characteristics of the billiard ball and some specified different initial conditions, the billiard ball would have knocked someone out.

As in the *PLIC* sense, we fix the characteristics of the billiard ball by making invariant its intrinsic properties (such as its size, weight, shape and hardness). As before, we also fix the laws of physics. Otherwise, we would be talking about what is logically possible and not just what is physically possible. In contrast with the *PLIC* sense, however, our conditional now ranges over different initial conditions while keeping the context of the situation the same. Using the terminology introduced in the previous section, the conditional ranges over different external circumstances within the given scenario. Our scenario consists in John playing billiards at his local hotel and sending the billiard ball flying towards the bar at high speed and at head height. To determine whether the billiard ball could have harmed someone, we ask: Are there any external circumstances that would have led to patrons being knocked out? Well, if Rudy had stood two metres to the left of where he was standing or Jane had stood one metre to her right, they would have been knocked out by the flying billiard ball. So, in this case, we had good reason to say that the billiard ball could have knocked someone out.

In their critiques of compatibilists, libertarians and hard determinists ignore this second sense, the *PLUC* sense, of 'could have done otherwise' when we speak of free human acts. By proposing that we are exclusively using the *PLIC* sense when we assert our free will, libertarians are forced to construct an untenable metaphysical story about how intentional human acts break the laws of physics. Scientifically-minded hard determinists, in also acquiescing to the *PLIC* sense, are thus compelled to follow where their reasoning takes them and declare the counterintuitive conclusion that we are not free to act. Once we recognize the *PLUC* sense of 'could have done otherwise' and how employing it in ordinary discourse does not imbue billiard balls and the like with contra-causality, we are in a position to do the same with talk of free agents. By applying the right semantic analysis, it becomes clear how human beings can act freely without contravening the laws of physics.

I want now to apply the *PLUC* sense of 'could have done otherwise' to a case involving a free act. The basic schema I am suggesting here is as follows. An agent could have done otherwise if:

Given the agent's character, the agent would have done otherwise in this situation if certain external circumstances were different.

Consider the Officer from the scenario pictured in the previous section waiting in line in the police canteen the morning of the shooting. Suppose that he likes both tea and coffee and has drunk them on different occasions. That morning in the canteen, say he chooses coffee at the dispenser to drink with his breakfast. Being a free act, he could have selected tea. On a compatibilist analysis employing the *PLUC* sense, his selecting coffee freely means that, given the Officer's character, the Officer would have done otherwise in this situation if certain external circumstances were different.

As per my previous analysis of unfree acts given in the previous section, what I mean by 'different circumstances' is possible circumstances that are external to the agent's character and that both preserve the opportunity for the agent to perform the act and the



situational context of the question at hand. In this case, if the canteen was out of coffee that morning, then the Officer would not have had the opportunity to have selected coffee and the question of his choosing coffee freely or not would not have arisen.

Given that the Officer had the opportunity to select coffee, were there any external circumstances that would have led the Officer to choose tea instead? Well, he would have selected tea with his morning breakfast if it had been offered at a discount price or if he had had coffee for the last five days or the canteen was trialling a new brand of tea leaves. We do not even need to know the exact circumstances that would have led to the Officer choosing tea that morning. Our assent to the counterfactual conditional only requires that we have reasonable confidence that given the Officer's character trait of liking tea and coffee, there are some circumstances in which he will choose tea.

Note how in employing a counterfactual conditional analysis using the *PLUC* sense, the concept of 'characteristics' of an inanimate object is commensurable with the concept of 'character' of a human agent. The size, weight, shape and hardness of the billiard ball that give it the ability to knock out a patron is commensurable with the police officer's beverage preferences, price sensitivity and propensity for experimentation that give him the ability to choose tea over coffee.

It should be noted at this point that the truth of the compatibilist's thesis in general and of my 4C theory of free acts in particular does not necessarily require that a free act is one for which the agent could have done otherwise. Some compatibilists have resolutely rejected this latter thesis. Frankfurt [1969], for example, has constructed counterexamples purportedly demonstrating the possibility of agents acting freely and for which the agent is morally responsible and yet could not have acted other than they did. In my Allan [2016], I try to show how the type of counterfactual conditional analysis offered in this essay survives Frankfurt counterexamples.

The distinction I drew above between *PLIC* and *PLUC* senses makes it clear that the type of analysis using counterfactual conditionals that I am proposing here is not some kind of special technical or philosophical translation of what ordinary folk mean by 'had no choice' and 'could have done otherwise'. I'm suggesting that this is how common folk understand these phrases when talking about free and unfree acts. This analysis is a more natural rendering of 'free will' talk compared with the burdensome metaphysical presuppositions of both the hard determinist and libertarian analysis.

That the *PLUC* sense of 'could have done otherwise' is the most natural sense is further supported by a review of how people use the term in common discourse when referring to inanimate things and processes. Consider this commonplace observation by Lakhani [2014]: 'Hurricane Katrina could have killed more people than the Galveston hurricane did.' According to Lakhani, Katrina didn't kill more people because it struck considerably later when the resident population was much less. Given Katrina's characteristics, it would have killed more people if it had struck at a different time, in different external circumstances. Examples like this abound in the common literature. The same mundane counterfactual conditional sense in terms of fixed characteristics and different circumstances is given to the number of people that could have been killed by a falling branch [Kisiel 2013], the number of animals that could have been saved from a flood

[Barber 2016] and the number of people that could have been killed by a bushfire [CBS News 2015].

As these ordinary-language examples illustrate, the kind of analysis using counterfactuals that helps us make sense of the application of the modal verb, 'could have', to inanimate objects likewise helps us make sense when applied to human agents. In either case, we have no need to import additional metaphysical presumptions about uncaused causes to carry through the analysis.

### 7.3 Difficult Cases

There are circumstances in which the Officer described in the scenario in §7.1 above does not freely select coffee with his breakfast that morning. He may have developed a strong addiction to caffeine, which then compels him to drink coffee exclusively. The 4C compatibilist theory of what makes an act freely chosen needs to account in a principled way for why this act is not free. Recall that on this theory, there are four requirements for an act to be freely chosen: (1) the act must not feel compelled; (2) the act must not be controlled by a third party; (3) the act results from the agent's character; and (4) the agent is cognitively capable of reasoning about the act. (See §4 above, where I introduce the 4C theory.)

If the Officer's addiction to caffeine is sufficiently strong, his choosing coffee fails requirements (1) and (3) at the least and so cannot be considered an act of free choice. His complaint that he 'had no choice' in choosing coffee is apt and we can interpret his failure in not being able to choose otherwise using the same counterfactual conditional analysis as before. That is, given the Officer's character (liking both tea and coffee), the Officer would not have done otherwise in this situation (addiction plus canteen breakfast), even if the external circumstances were different. In this case of addiction, the Officer feels psychologically compelled to select coffee and thereby violates requirement (1) for a free act.

There are situations in which the Officer may have felt no compulsion to choose coffee and yet the act is considered by most people as unfree. I want now to turn our attention to these kinds of cases. Consider a situation in which the Officer instead underwent an unusual operation the night before. Unbeknownst to the Officer, a neurosurgeon implanted a microchip with remote communication capabilities in his brain that allowed the neurosurgeon to control the Officer's desires from a remote location. Moments prior to choosing his beverage the next morning to enjoy with his breakfast, the neurosurgeon remotely fixes the neuronal firings in the Officer's brain so that he desires and chooses coffee.

The Officer feels no compulsion to choose coffee, so requirement (1) for a free act is satisfied. He may even be able to adduce reasons for his choosing coffee, so requirement (4) may be satisfied. If the neurosurgeon's tinkering is not too dramatic, we may even be able to say that the Officer's choice is consonant with and a consequence of his character. So, requirement (3) may be satisfied. However, as the Officer's intention to choose coffee is directly controlled by the neurosurgeon from afar, his selecting coffee fails requirement (2).

In this case, then, it is apt to say that the Officer could not have done otherwise. The counterfactual conditional analysis of his inability to do other than he did is this: given the Officer's character (liking both tea and coffee), the Officer would not have done otherwise in this situation (remote control plus canteen breakfast) even if the external circumstances were different. It is also reasonable to say that the Officer 'had no choice' in this circumstance even though to the uninformed observer it seems he did have a choice. In this case, the agent doing the choosing is the neurosurgeon, with the unsuspecting Officer simply acting as his proxy.

Consider one final situation in which the Officer does not freely select coffee with his breakfast that morning. Imagine that the Officer has a family history of mental illness. While waiting in the canteen queue that morning, he experienced, for the first time, a mild psychotic episode in which he believed that God was commanding him to select coffee in order to fulfil some divine purpose. As he feels psychologically compelled to obey this divine stricture, requirement (1) is not satisfied. Requirement (3) is also not satisfied, as his actions are out of keeping with his character. Individuals experiencing a psychotic episode are unable to reason clearly about their actions, so requirement (4) also fails. The counterfactual conditional analysis of the fact that he could not have done other than to select coffee follows the same schemata as per the previous examples.

In this section, I considered how we can make sense of an agent acting freely within a deterministic system by employing an ordinary-language counterfactual conditional analysis of 'could have done otherwise'. I then subjected this analysis and the 4C compatibilist theory to difficult cases in which people act unfreely and which severely test our notions of free will and human agency. If the counterfactual conditional analysis and the 4C theory are to be vindicated, they need to be able to account in a principled, consistent and coherent way how the agents of these acts could not have acted otherwise.

Although I think this task has been achieved, I am not claiming an easy victory. Many of these cases involving mental illness, addiction, brain microchips and genetic manipulation generate much discussion and disagreement amongst philosophers, ethicists, judges and clinicians. However, I think the analytical tools advanced in this essay provide a framework for rational discussion about cases located within these grey areas. Perhaps to their credit, the ideas encapsulated in the 4C theory are not new. They have served as the backbone in the theory and practice of jurisprudence for well on a century.

## 8. Determinism, Free Will and Moral Responsibility

In this final section, I want to explore briefly how living in a deterministic world in which all of our behaviours are, in principle, predictable is consistent with the notion of moral responsibility and with our common practice of praise and blame. I will conclude with some remarks on the role that our concept of free will plays in this deterministic scheme.

To begin with, knowing that our fellow citizens do what they do as a result of an extremely complex interplay between their genes, their physical environment, socialization and just plain luck of where and when they were born ought to instil within us a greater understanding of their social limitations and achievements. When it comes to punishing people who do to us and the broader community wrong, such a global perspective ought to make us less prone to want to inflict pain on wrongdoers simply for the sake of revenge. This more enlightened perspective should extend even to people outside of our own communities; that is, to all of the fellow inhabitants of our planet Earth.

Such a deterministic view of human behaviour naturally works against a retributive approach to punishment. I am pleased to see that in secular democracies at least, the developing scientific view of the world that has given us a much greater understanding of psychiatry, psychology, neuroscience and the social determinants of behaviour has bore witness to a much more humane approach to criminal justice over the last one hundred years or so. In scientifically informed nations, other approaches to crime and punishment have increasingly taken the place of earlier retributive approaches. These include the idea of instituting punishment to deter others, incarceration to prevent further criminal acts, voluntary and mandatory treatment regimes that aim to rehabilitate criminals and restitution programs to recompense victims of crime.

We all have an interest in maintaining our common moral and legal systems. These systems work to ensure the safety and security of our persons, our family and others we care about. They also protect our livelihood, allowing us to go about our business unimpeded and to pursue endeavours that enrich our lives and those of our loved ones. This is true of all people and irrespective of whether the world is only partly or wholly deterministic. Whether we are supremely complex thinking and feeling wet robots or not, we possess these interests in our own welfare and the welfare of others and in promoting the moral and legal systems that guarantee them.

In a deterministic world, our concept of free will has an integral part to play in our moral and legal institutions. Of course, this is not 'free will' in the contra-causal sense. In fact, I find it hard to make sense of the notion of moral and legal responsibility in a world where particular thoughts, desires and intentions can pop into existence without any grounding in a person's character or past.

There are good, positive reasons for thinking that indeterminism cannot be the basis for moral responsibility. Take a morally praiseworthy or blameworthy act, such as lending money to a criminal. Suppose Mary lends some money to Joel. It would not be out of the ordinary or irrational to express the counterfactual conditional: 'If Mary had known Joel is using borrowed money for criminal purposes, she would not have loaned money to him.' However, if we accept the libertarian's thesis that free acts are those that have no sufficient

psychological, neurophysiological or divine cause, then how are we to know whether the counterfactual conditional is true or false? What sense could we make of the necessity of the consequent ('Mary does not lend money to Joel') given the antecedent ('Mary knows Joel is using borrowed money for criminal purposes')?

Counterfactuals such as these only make sense and allow evidence to be adduced for them being true or false if moral agents act in law-like ways. (In my Allan [2015], I discuss in more detail how, if we assume indeterminism, all counterfactual conditionals about free human choices are either without a truth value or are false.) In moral reasoning, saying how moral agents would have acted in different circumstances figures highly in determining guilt and innocence. Given the part counterfactual reasoning plays in our moral and legal judgements, far from moral reasoning being antithetical to deterministic thinking, it seems to require it.

In the specific compatibilist sense I have been advancing in this essay, the concept of 'free will' has a crucial role to play in identifying those intentions and actions that are morally praiseworthy and blameworthy and those that are legally excusable and culpable. In the moral sphere, placing importance on free acts starts from a very early age. We keep a look out for behaviours from our children that harm others, as these instances provide opportunities for character building through moral guidance and appropriate praise and blame. Every capable parent recognizes that it's of no use, and even counterproductive, to praise or blame a child for an act that they did not choose freely. When a child has been coerced into hitting another or that child falls on another child while doped up on medicine, the child's parents avoid admonishing her. Parents pay particular attention to their child's intentions, even the ones that do not lead to action, as having the right intentions builds character and leads to admirable behaviour. Knowing whether the child's will is formed freely or not in a particular instance is crucial in helping parents decide whether praise or blame in that instance will further their child's moral education.

Just as in the moral sphere competent parents do not praise or blame children for unfree acts, communities do likewise in the realm of criminal justice. Punishing and rewarding citizens for acts that were not freely chosen is irrational and counterproductive. For example, punishing mentally ill people for actions over which they had no control serves no useful purpose. As lawyer Julie Grachek explains, 'if the mental illness has caused loss of free will, the offender has lost his ability to freely choose whether or not to recommit the offense, and he is therefore undeterrable through punishment' [Grachek 2006: 1482].

The concepts of 'free will' and 'free act' serve as key indicators in both the moral and legal spheres for whether a person is responsible for their action. We express this typically in the axiom that the exercise of a person's free will is a necessary precondition for holding that person responsible for their act. In other words, a person must have acted freely before we deem their act morally and legally praiseworthy or blameworthy. During the course of this essay, I offered a number of examples in which a person's free will is compromised, either through coercion, manipulation, addiction or mental illness, and how that encumbrance gives us reason to excuse or mitigate that person's culpability.

## 9. Conclusion

Hard determinists and libertarians agree that if determinism is true, we ought to give up 'free will' talk. These incompatibilists argue that if all human thoughts and actions are the result of sufficient prior causes, then we could not have acted otherwise than what we actually did. In this essay, I tried to show that this line of reasoning is mistaken. To the extent that incompatibilists are referring to the way we ordinarily speak about free will, I argued that their analysis is misconstrued. By tracing the evolution of 'free will' language through its sixteenth century origins to its use in ordinary discourse today, I showed that for the person in the street, lack of restriction and not contra-causality is the central meaning of 'free will'.

Through examining discourse about more complex situations in which people are manipulated and suffer addiction and mental illness, I distilled the key factors that are commonly thought to restrict a person's free will. With these limiting factors, I forged them into the '4C theory' of compatibilism: Compulsion, Control by a third party, Character loss and Cognitive inability to reason. A key feature of this more comprehensive compatibilist view is that it reflects and encapsulates how modern thinking in the areas of medicine and jurisprudence has led to the further development of our notion of 'free will'. These four 'free will' limiting factors are now recognized in medical and legal practice today and are mainstay in courtrooms throughout the developed world.

At the pointy end of the stick, I also argued that a little bit of philosophy and science can be dangerous in the hands of incompatibilists. In paying insufficient heed to how ordinary folk engage in 'free will' talk, incompatibilists superimpose their own metaphysical presuppositions onto the meaning of 'free will'. This unconscious metaphysical overlay on common meanings lead them to suppose erroneously that ordinary folk 'intuit' that their wills are contra-causal. I ended this essay with a reflection on the linkage between our concept of 'free will' and our moral practices. I tried to show that far from our moral discourse presuming we have contra-causal free will, it is in reality underpinned by a determinist view of human behaviour.<sup>10</sup>

---

<sup>10</sup>I am grateful to the many people who submitted comments and criticism of earlier versions of this essay. I am especially indebted to Pierre-Normand Houle for his detailed commentary and for our long and fruitful conversations on this subject. I remain wholly responsible for any errors and omissions in the published version.

## References

- Allan, Leslie 2015. [Plantinga's Free Will Defence: Critical Note](http://www.RationalRealm.com/philosophy/metaphysics/plantinga-free-will-defence.html), URL = <<http://www.RationalRealm.com/philosophy/metaphysics/plantinga-free-will-defence.html>>.
- Allan, Leslie 2016. [Frankfurt Cases and 'Could Have Done Otherwise'](http://www.RationalRealm.com/philosophy/metaphysics/frankfurt-cases.html), URL = <<http://www.RationalRealm.com/philosophy/metaphysics/frankfurt-cases.html>>.
- Barber, Kayleigh 2016. Animals That Nearly Drowned in Zoo Flood Back in Captivity, Just in New Cages, *One Green Planet*, URL = <<http://www.onegreenplanet.org/animalsandnature/animals-that-nearly-drowned-in-zoo-flood-back-in-captivity-just-in-new-cages/>> (Published: January 13, 2016).
- Before It's News 2012. How Obama Is Manipulating People Utilizing Mass Hypnosis Techniques, *Before It's News*, URL = <<http://beforeitsnews.com/obama/2012/12/how-obama-is-manipulating-people-utilizing-mass-hypnosis-techniques-2446468.html>> (Published: December 19, 2012).
- Berofsky, Bernard, ed., 1955. *Free Will and Determinism*, New York: Harper and Row.
- Brown, Andrew M 2016. Obesity: Addiction? Or Free Will?, *The Telegraph*, URL = <<http://blogs.telegraph.co.uk/news/andrewmcbrown/100031840/obesity-addiction-or-free-will/>> (Published: April 5, 2016).
- Campbell, Charles 1967. In Defence of Free Will, *In Defence of Free Will*, London: Allen and Unwin.
- Catholic Education Daily 2015. Hypnotism Poses Real Dangers, despite Popularity on College Campuses, *Catholic Education Daily*, URL = <<http://www.cardinalnewmansociety.org/CatholicEducationDaily/DetailsPage/tabid/102/ArticleID/3970/Hypnotism-Poses-Real-Dangers-despite-Popularity-on-College-Campuses.aspx>> (Published: February 11, 2015).
- CBS News 2015. More than 1,000 Homes Destroyed by 2 California Fires, *CBS News*, URL = <<http://www.cbsnews.com/news/more-than-1000-homes-destroyed-by-2-california-fires/>> (Published: September 19, 2015).
- Colorado v. Connelly* 1986. URL = <[http://www1.law.umkc.edu/suni/wrongful\\_convictions/colorado\\_v\\_Connelly.htm](http://www1.law.umkc.edu/suni/wrongful_convictions/colorado_v_Connelly.htm)>.
- Cummins, Denise 2014. The Myth and Reality of Free Will: The Case of Addiction, *Psychology Today*, URL = <<https://www.psychologytoday.com/blog/good-thinking/201402/the-myth-and-reality-free-will-the-case-addiction>> (Published: February 9, 2014).
- Debate.org 2016. When Under an Addiction, Does the Addict Still Have Free Will?, *Debate.org*, URL = <<http://www.debate.org/opinions/when-under-an-addiction-does-the-addict-still-have-free-will>> (Retrieved: April 5, 2016).



- De Wolf, Aschwin 2011. Neural Cryobiology and the Legal Recognition of Cryonics, *Institute for Evidence-Based Cryonics*, URL = <http://www.evidencebasedcryonics.org/2011/05/16/neural-cryobiology-and-the-legal-recognition-of-cryonics/> (Published: May 16, 2011).
- Dickens, Charles 1843. *The Life and Adventures of Nicholas Nickleby*, vol. 2, Leipzig: Bernhard Tauchnitz.
- Dowling, James 2016. Police Officer Thought Terror Teen Numan Haider Would Behead Him, Inquest Hears, *Herald Sun*, URL = <http://www.heraldsun.com.au/news/law-order/police-officer-thought-terror-teen-numan-haider-would-behead-him-inquest-hears/news-story/dc59c3ca81f98f0b50300c7bcb974b84> (Published: March 17, 2016).
- Eichel, Steve K. D. 2004. Commonwealth of Virginia v. Lee Boyd Malvo, (aka “Beltway Snipers”), URL = [http://www.dreichel.com/malvo\\_trial.htm](http://www.dreichel.com/malvo_trial.htm) (Published: January 22, 2004).
- Emmanuel, Chidi and Sahar Moussa 2015. Arranged Marriage Yes I Do, No I Don’t, *Kuwait Times* online, URL = <http://news.kuwaittimes.net/arranged-marriage-yes-i-do-no-i-dont/> (Published: April 9, 2015).
- Frankfurt, Harry 1969. Alternate Possibilities and Moral Responsibility, *Journal of Philosophy* 66: 829–39.
- Glannon, Walter 2014. The Free-Will Fix, *Aeon*, URL = <https://aeon.co/essays/how-new-brain-implants-can-restore-and-enhance-free-will> (Published: October 1, 2014).
- Grachek, Julie E. 2006. The Insanity Defense in the Twenty-First Century: How Recent United States Supreme Court Case Law Can Improve the System, URL = <http://www.repository.law.indiana.edu/ilj/vol81/iss4/14> *Indiana Law Journal* 81/4: Article 14.
- Hardin, Herschel 1993. Uncivil Liberties, *Mental Illness Policy Org*, URL = <http://mentalillnesspolicy.org/media/bestmedia/uncivilliberties.html> and reprinted with permission from *The Vancouver Sun* (Published: July 22, 1993).
- Hodges, Dave 2015. Government Sponsored Mind Control Will Soon Eliminate All Free Will, *Fourwinds10.com*, URL = [http://www.fourwinds10.net/siterun\\_data/health/mind\\_control/news.php?q=1444322184](http://www.fourwinds10.net/siterun_data/health/mind_control/news.php?q=1444322184) (Published: October 7, 2015).
- Hoefer, Carl 2016. Causal Determinism, *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), ed. Edward N. Zalta, URL = <http://plato.stanford.edu/archives/spr2016/entries/determinism-causal/>.
- Jeffery, Jason 2016. Brain Zapping Electronic Mind Control: Part 1, URL = <http://www.whale.to/b/jeffery.html> (Retrieved: April 2, 2016).

- Johnston, Terilyn Lee 2002. Neurology and Mind Control: A Beneficial Technology or the Opening of Pandora's Box?, *Science, Technology, and the Pursuit of Truth*, URL = <[http://people.bethel.edu/~kisrob/hon301k/project/Final/mind\\_control.htm](http://people.bethel.edu/~kisrob/hon301k/project/Final/mind_control.htm)> (Published: December 15, 2002).
- Kisiel, Ryan 2013. Did Kew Gardens Staff Fail to Properly Prune Trees before Death of Accounts Manager, *Daily Mail*, URL = <<http://www.dailymail.co.uk/news/article-2440017/Did-Kew-Gardens-staff-fail-prop-erly-prune-trees-death-accounts-manager-31-killed-falling-branch.html>> (Published: October 2, 2013).
- Lageman, Thessa 2016. The Dark Side of Dubai: Human Trafficking and Prostitution, *The Age*, URL = <<http://digitaledition.theage.com.au/Olive/ODE/online/LandingPage/LandingPage.aspx?href=QUdFLzlwMTYvMDEvMjE.&pageno=MTQ.&entity=QXlwMTQwMg.&view=ZW50aXR5>> (Published: January 21, 2016, The World section, p. 14).
- Lakhani, Aiman 2014. Hurricanes, *Prezi*, URL = <<https://prezi.com/tmivsnzs7b2i/hurricanes-by-aiman-lakhani/> 19 November 2014> (Published: November 19, 2014).
- Law Library 2016. Guilty Plea: Accepting the Plea - The Elements of Guilty Pleas, *Law Library*, URL = <<http://law.jrank.org/pages/1277/Guilty-Plea-Accepting-Plea-elements-guilty-pleas.html>> (Retrieved: April 3, 2016).
- Moyers, Bill 1998. An Interview with Alan I. Leshner, Ph.D., *Moyers on Addiction: Close to Home*, URL = <<http://www.thirteen.org/closetohome/science/html/leshner.html>> (Retrieved: April 3, 2016); URL = <[http://www.thetherapist.com/PBS\\_Article\\_02.html](http://www.thetherapist.com/PBS_Article_02.html)> (Retrieved: May 9, 2016).
- McKenna, Michael and D. Justin Coates 2015. Compatibilism, *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), ed. Edward N. Zalta, URL = <<http://plato.stanford.edu/archives/sum2015/entries/compatibilism/>>.
- Nemade, Rashmi and Mark Dombeck 2009. Symptoms of Schizoid Personality Disorder, *MentalHelp.net*, URL = <<https://www.mentalhelp.net/articles/symptoms-of-schizoid-personality-disorder/>> (Published: August 7, 2009).
- Ninemsn 2016. Police Officer Feared He'd Be Beheaded by Radicalised Teen Numan Haider, *Ninemsn*, URL = <<http://www.9news.com.au/national/2016/03/17/18/50/police-officer-feared-hed-be-beheaded-by-radicalised-teen-numan-haider>> (Published: March 17, 2016).
- O'Connor, Timothy, Free Will, *The Stanford Encyclopedia of Philosophy* (Fall 2014 Edition), ed. Edward N. Zalta, URL = <<http://plato.stanford.edu/archives/fall2014/entries/freewill/>>.

- Oliphant, Roland 2016. Nanny Beheaded Four-Year-Old 'In Revenge for Russian Bombing of Syria', *The Telegraph*, URL = <http://www.telegraph.co.uk/news/worldnews/europe/russia/12181937/Nanny-be-headed-four-year-old-in-revenge-for-Russian-bombing-of-Syria.html> (Published: March 3, 2016).
- Robinson, Paul H. 1982. Criminal Law Defenses: A Systematic Analysis, *Columbia Law Review*, URL = <https://www.law.upenn.edu/fac/phrobins/defenses.html>.
- Sky News 2013. Peru Drugs Bust: Pair 'Forced at Gunpoint', *Sky News*, URL = <http://news.sky.com/story/1127895/peru-drugs-bust-pair-forced-at-gunpoint> (Published: August 14, 2013).
- SBS 2014. Comment: I Was Brainwashed to Believe Kim Jong-Il Was a God, *SBS*, URL = <http://www.sbs.com.au/news/insight/article/2014/04/07/comment-i-was-brainwashed-believe-kim-jong-il-was-god> (Published: April 14, 2014).
- West, Nicholas 2013. 7 Future Methods of Mind Control, *ActivistPost.com*, URL = <http://www.activistpost.com/2013/06/7-future-methods-of-mind-control.html> (Published: June 2, 2013).
- Wikipedia 2016a. Medical Definition of Death, *Wikipedia*, URL = [https://en.wikipedia.org/wiki/Medical\\_definition\\_of\\_death](https://en.wikipedia.org/wiki/Medical_definition_of_death) (Retrieved: April 9, 2016).
- Wikipedia 2016b. Cryonics, *Wikipedia*, URL = <https://en.wikipedia.org/wiki/Cryonics> (Retrieved: April 9, 2016).
- Wright, Chris 2013. Do Addicts Have Free Will, *Alternet*, URL = <http://www.alternet.org/do-addicts-have-free-will> (Published: November 30, 2013).