

THE CORRELATES AND CONSEQUENCES  
OF BELIEVING IN FREE WILL

BY  
MELISSA SNATER

A thesis  
submitted to the Victoria University of Wellington  
in fulfilment of the requirements for the degree of  
Master of Arts

Victoria University of Wellington  
2019

## TABLE OF CONTENTS

|   |    |
|---|----|
| Abstract .....  | i  |
| Acknowledgements .....  | ii |
| 1. INTRODUCTION .....   | 1  |
| 1.1 Nondisclosure of scientific doubt about free will .....               | 3  |
| 1.2 Illusionism.....  | 5  |
| 1.3 Disillusionism .....  | 6  |
| 1.4 Free will and moral responsibility.....                               | 8  |
| 1.5 Summary .....   | 12 |
| 2. THE EMPIRICAL STUDIES .....  | 13 |
| 2.1 Evaluating free will beliefs and moral behaviour .....                | 14 |
| 2.2 Determinism, fatalism, and the problem of priming.....                | 17 |
| 2.3 Failure to replicate and unvalidated measurement tools.....           | 21 |
| 2.4 An upside to desert scepticism.....                                   | 22 |
| 2.5 Two commentaries on the social implications of belief in free will .. | 26 |
| 2.6 A mechanistic view of human nature reduces retribution.....           | 28 |
| 2.7 Are people natural compatibilists or incompatibilists? .....          | 30 |
| 2.8 Summary .....   | 32 |
| 3. THE FREE WILL AND DETERMINISM SCALE .....                              | 34 |
| 3.1 The free will inventory .....   | 39 |
| 3.2 Scale development .....   | 42 |
| 3.3 Predictions.....  | 43 |
| 4. METHOD.....  | 44 |
| 4.1 Participants.....   | 44 |
| 4.2 Materials and procedure.....  | 44 |
| 5. RESULTS .....  | 48 |
| 5.1 Discussion .....  | 50 |
| 5.2 Correlations and construct validity .....                             | 51 |
| 5.3 Discussion .....  | 52 |

|  |    |
|--|----|
| 5.4 The effects of real-world scepticism ..... | 53 |
| 5.5 Discussion .....                           | 56 |
| 5.6 General discussion .....                   | 57 |
| 6. CONCLUSION .....                            | 61 |
| Bibliography.....                              | 62 |

## ABSTRACT

Research has indicated that weakening people's belief in free will may likewise weaken their belief in moral responsibility and potentially license them to morally transgress. Recent studies in social psychology suggest that diminished belief in free will is associated with a range of anti-social or otherwise negative outcomes. For example, cheating, unjustified aggression, and less prosocial helping behaviour. In response to these findings, *illusionist* philosophers have recommended that even if scientists somehow conclusively showed that free will does not exist it might nevertheless be necessary to foster widespread belief as a useful fiction. In the opposing camp, free will *disillusionists* maintain that belief in free will has a dark side that we would be better off without. The problem they say, is the close connection between free will and the belief that people *justly deserve* what they get. So rather than having the instrumental benefits that illusionists claim, belief in free will is too often taken to justify treating people in severe and demeaning ways. Who then is correct? I report empirical results comparing the beliefs and attitudes of free will sceptics and people naïve to the debate. Results are consistent with the claims of disillusionists. Free will sceptics are more compassionate and are less likely to believe in just deserts and harbour retributive attitudes.

## Acknowledgements

First and foremost, I want to thank my supervisor Dr Justin Sytsma for the wealth of feedback, suggestions and direction. His efforts were hugely instrumental for the development of this thesis and for the improvement of my skills as a researcher and writer. I would also like to thank Victoria University of Wellington for funding this project.

## 1. INTRODUCTION

Few concepts are more ingrained in the collective Western psyche than free will. While there is considerable disagreement among philosophers about how the term free will should be best understood. The belief that we can (at least sometimes) make decisions and perform actions that are free from the physical constraints of our biology and environment is a pervasive Western assumption. Yet, not only is there no evidential support for such a capacity, the latest research about how the brain functions appears to point in the opposite direction. New discoveries increasingly demonstrate the many ways in which humans are subject to influences well beyond the control of the conscious self. Given that our Western sensibilities generally take free will to be the linchpin upon which all moral obligations depend (Kant, 1785), it is unsurprising that many people feel threatened by the prospect that, despite its intuitive appeal, free will may be an illusion. In the past, much of the concern surrounding free will scepticism was the more or less speculative worry that widespread acceptance of such scepticism would lead to the unravelling of our moral fabric. Recently this worry has been given empirical support. New findings from social psychology indicate that diminished belief in free will is associated with reduced prosocial tendencies and increased anti-social tendencies. In response to these findings, many academics have expressed concern that widespread doubt about the existence of free will may have a deleterious effect on individual behaviour, interpersonal relationships and the political and social institutions that rely on the assumption of it. Some have even gone so far as to recommend insulating the public from information that calls into question the existence of free will. Given how strongly this message has been resonating both within the academic community and more recently the public domain, a detailed examination of whether or not society would benefit by giving-up the belief in free will is necessary and timely.

In this thesis, I set out to critically analyse some of the original studies that purport to show negative consequences for free will scepticism, I conclude that their results are not sufficient to justify such extreme measures. Using empirical methods from social psychology, I proceed to investigate the real-world consequences of believing in free will, not believing in free will, or believing in it to varying degrees under different descriptions. Drawing on the arguments of free will sceptics, previous research, and my own empirical

investigations I show that belief in free will actually has a dark side we would be better off without. I present evidence that lends strong support to the philosophical view that belief in free will is closely linked to the belief that individuals *justly deserve* what they get, and that this belief is frequently called in aid to justify retributive attitudes that are highly damaging to society.

It is worth stating at the onset that my concern in this thesis is not with whether or not scepticism is true. My concern here is primarily with determining whether the available empirical evidence that suggests adopting a sceptical world-view would have an adverse effect on the moral behaviour of ordinary people.

The current study seeks to fill a number of gaps in the extant literature by developing and testing two novel approaches to free will research. In the process I make three important new contributions to the free will literature. (1) Where previous correlational studies examining the links between belief in free will and various other beliefs have assumed a unified model of free will belief, I have developed and validated a new measurement tool on two dimensions that demonstrates belief in libertarian free will (*contra-causal agency*) and the endorsement of compatibilist views can be studied independently. (2) Operationalising free will belief on these two dimensions enabled me to find the long sought-after inverse relation between belief in contra-causal agency and belief in determinism. Thus, for the first time in psychometric work on free will lay incompatibilism is indicated. (3) *All* previous behavioural research on belief in free will has relied on priming techniques to induce disbelief. The current study is the first of its kind to examine the beliefs and attitudes of free will sceptics in the real world.

I proceed as follows. In Chapter 1, I provide a general summary of the motivation for the claims for and against public disclosure of scientific doubt about free will and present the arguments from *illusionism* and *disillusionism*. I then situate moral responsibility within the context of the free will debate and discuss the entailment of just-deserts and the justification for retributive punishment. In Chapter 2, I critically analyse the extent to which some of the previous studies suggest reduced belief in free will would negatively impact society, I then lay out the motivations for free will scepticism and discuss several supporting empirical studies. Finally, I discuss work in X-phi on the problem of natural (in)compatibilism. In Chapter 3, previous psychometric methodology and the rationale for a two-dimension model of free will belief is discussed and

predictions are laid out. In Chapter 4, the method is detailed. In Chapter 5, the results are reported and discussed. Chapter 6, Conclusions are drawn from findings.

## 1.1 NONDISCLOSURE OF SCIENTIFIC DOUBT ABOUT FREE WILL

Whilst most people believe in free will (Nahmias et al., 2005),<sup>1</sup> long standing arguments remain about what sort of free will can realistically exist alongside natural physical laws. While few people would deny that we regularly exercise self-control and make uncoerced choices, many scientists and philosophers take issue with the idea that humans can generate conscious decisions and actions that are not fully determined by prior events. According to this challenge *contra-causal free will* (free will that operates outside of natural laws and allows for causally untraceable actions) is incompatible with a scientific understanding of the world in which all events and actions, including human action, are the inevitable outcome of prior causes.<sup>2</sup> It therefore stands to reason that a growing number of theorists and researchers have become increasingly vocal about their scepticism concerning contra-causal free will, with many arguing it is an illusion that is growing less and less believable under accumulating evidence supporting a more mechanistic view of human behaviour (Bargh, 2008; Green & Cohen, 2011; Shariff et al., 2014; Wegner, 2002). Consequently, with articles espousing scientific doubt appearing in the *New York Times*, *The Atlantic*, and recently the *Scientific American* the free will debate has hit the mainstream and anti-free will sentiments are on the rise.

A common theme about the way free will is represented in the popular press is in terms of the recent scientific discoveries, especially in neuroscience, that threaten or undermine our intuitive sense of agency and moral responsibility. Chief among these is a contemporary picture of the brain, in which *all* brain activity is caused by other brain activity in a densely interconnected neuronal network that leaves no room for agency that

---

<sup>1</sup> Most of the of the experimental studies on lay belief in free will have been conducted in American samples, so “most people believe in free will” is not a universal claim and caution is needed in generalising to populations outside the U.S.

<sup>2</sup> Even though the standard interpretation of quantum mechanics has been taken by many to undermine the thesis of universal determinism, sceptical arguments generally hold that even if we allow some indeterminacy to exist at the micro level of our existence – the level studied by quantum mechanics – there would still likely remain determinism where it counts at the level of human action (Honderich, 1988, as cited in Caruso, 2013).



is “free” from the rules of biology (Eagleman, 2011). Also, the striking discovery that unconscious brain activity causally initiates action prior to conscious awareness of the intention to act (Libet et al., 1983; Brass, Demanet & Rigoni, 2013). Other studies, in cognitive and social psychology on *automaticity and situationism* have revealed just how susceptible our internal psychological processes are to the influence of events in our immediate environment with no conscious awareness of such (Bargh & Ferguson, 2000; Wilson, 2002). For many, these findings represent a serious threat to our ordinary understanding of ourselves as conscious, rational, morally responsible agents, since they show that the conscious mind exercises less control over our behaviour than we traditionally assumed. Reflecting these sentiments Dennis Overbye of the *New York Times* writes, “the death of free will, or its exposure as a convenient illusion, some worry, could wreak havoc on our sense of moral and legal responsibility ... it would mean that people are no more responsible for their actions than asteroids or planets, anything would go” (2007, p. F1).

Regardless of whether the scientific sceptics are well-motivated or misplaced about the putative death of free will, the idea is that their proclamations could still have an effect on people’s moral attitudes and behaviour. Free will sceptic Saul Smilansky (2000) maintains that not only do most people believe in their actual ability to transcend their circumstances, they have strong beliefs that libertarian free will is a condition for moral responsibility, which in turn provides the justification for reward and punishment. Smilansky believes that, if we are to destroy such beliefs the difficulties caused by the absence of moral desert will threaten common sense morality and cause acute psychological discomfort for many people.<sup>3</sup> Smilansky’s claims raise important philosophical questions about the psychological and societal consequences for society if free will scepticism is widely accepted. For example, if moral behaviour depends on belief in free will, and knowledge of determinism takes hold, will it leave us morally irresponsible? And if free will is an illusion what might happen to our institutions that are based on the assumption of it?

Even though philosophical debates about the metaphysical and moral status of free will have persisted for centuries without end, and without any clear consensus

---

<sup>3</sup> Smilanskys assertion that most people hold strong libertarian beliefs is speculative, whether or not lay-people generally have libertarian beliefs is an open question.

forming. Philosophical disagreement about the nature and existence of free will does not preclude the possibility of making progress on the sorts of empirical questions that are raised by Smilansky's concerns. While there was some pioneering work undertaken in the 1980s on the psychology of believing in free will (Viney, Waldman & Barchilon, 1982; Stroesner & Green, 1989), it was not until recently that a critical mass of researchers began to focus on what role, if any, believing in free will plays in our lives. To that end, the last decade has seen a large number of published studies in both social psychology and experimental philosophy with several suggesting that Smilansky's worries are, in fact, well founded. Notable findings have shown that inducing people to disbelieve in free will increases aggression, reduces helping, encourages cheating, and undermines moral behaviour. Other studies have found that diminished belief in free will makes people less creative, less likely to learn from their mistakes, and less grateful toward one another (Crescioni et al., 2016; Baumeister, Masicampo & DeWall, 2009; Vohs & Schooler, 2008; Shariff, Schooler & Vohs, 2008). These and other similar findings suggest it is not an overstatement to say that belief in free will plays a distinct role in motivating various positive traits, attitudes and behaviours. It has therefore been suggested that even though there is no empirical evidence to support free will there may be instrumental value in retaining the belief as a useful fiction. Some academics have even gone so far as to recommend deliberately shielding the public from information that calls the existence of free will into question as we will see.

## 1.2 ILLUSIONISM

Smilansky, is the frontrunner in a recent philosophical tradition known as *free will illusionism*. Like most philosophers he is convinced contra-causal free will does not exist but asserts that because illusory beliefs concerning contra-causal free will underpin folk assumptions about moral responsibility, to prevent people from embracing “an unprincipled nihilism” (2000, p. 189) free will beliefs should be preserved as much as possible. As Smilansky says, “To put it bluntly: people as a rule ought not to be fully aware of the ultimate inevitability of what they have done, for this will affect the way they hold themselves morally responsible” (2001, p. 85) Furthermore, he believes that scientists are unaware of the effects these ideas can have, and that promoting determinism

is complacent and dangerous (as cited in Cave, 2016). Thus, according to illusionism, people should be allowed to keep their positive illusion of free will, and those of us who have already been disenchanted simply ought to keep it to ourselves.

Similar illusionist views are found among psychologists. Kathleen Vohs and Johnathon Schooler conclude from their findings, “Although the concept of free will remains scientifically in question results point to a significant value in believing free will exists. If exposure to deterministic messages increases the likelihood of unethical action, then identifying approaches for insulating the public against this danger becomes imperative” (Vohs & Schooler, 2008, p. 54). Hence, Smilansky, Vohs and other illusionists maintain that even if scientists and philosophers were to conclusively show that free will does not exist, it might be nevertheless beneficial to foster widespread belief in free will to keep society from fraying at the seams (Baumeister & Monroe, 2014; Crescioni et al., 2016).

### 1.3 DISILLUSIONISM

On the other side of the fence, while most of the empirical work to date has focused on the potential *upside* to believing in free will, there is also a small but growing body of evidence supporting the idea that free will scepticism may have certain *positive* effects. New findings from psychology show reduced belief in free will has a significant impact on how people attribute moral responsibility, which in turn reduces attitudes associated with intolerance, victim blaming, and retribution (Carey & Paulhus, 2013; Nadelhoffer & Tocchetto, 2013; Shariff et al, 2014). Other studies have shown that stronger belief in free will is associated with increased support for economic inequality, and weaker belief is associated with a type of humility (Mercier et al., 2018; Earp et al., 2018). Such findings indicate that rather than bringing about negative consequences, widespread scepticism could potentially free us from a number of harmful tendencies, beliefs and attitudes.

In direct contrast to Smilanskys illusionism, Thomas Nadelhoffer defends *free will disillusionism*: “the view that to the extent that folk intuitions and beliefs about the nature of human cognition and moral responsibility are mistaken, philosophers and psychologists ought to do their part to educate the public – especially when their mistaken beliefs arguably fuel a number of unhealthy emotions and attitudes such as revenge,

hatred, intolerance, lack of empathy, etc.” (2011, p. 184). According to Nadelhoffer, “humanity must get beyond this maladaptive suite of emotions if we are to survive.” He adds that, “To the extent that future developments in the sciences of the mind can bring us one step closer to that goal – by giving us a newfound appreciation for the limits of human cognition and agency – I welcome them with open arms” (2011, p. 184). Nadelhoffer’s policy of disillusionism is also operative in the optimistic scepticisms of Derk Pereboom and Bruce Waller.

Pereboom (2001) believes that living without belief in free will would not be as detrimental to society as many people think. For instance, the possibility of finding meaning in life and sustaining good interpersonal relationships would not be jeopardised. And although certain reactive attitudes such as resentment and indignation would be undermined, he thinks these sorts of desert-based judgments are unnecessary for healthy well-functioning personal relationships in the first place. In fact, Pereboom claims that dispensing with the belief in free will is more likely to improve our well-being and relationships with others because it would tend to temper destructive forms of moral anger. Moreover, at the institutional level, although retribution and severe punishment, such as the death penalty would be ruled out, he notes that consequentially motivated sanctions like incapacitation, deterrence, and rehabilitation would still be justified.

Bruce Waller also makes a strong case for a world without belief in moral responsibility. In his book *Against Moral Responsibility* (2011), he cites many instances in which moral responsibility practices are counterproductive from a practical and humanitarian standpoint – most notably because they often encourage punitive excess in criminal justice and perpetuate social and economic inequalities. Waller suggests there would be great advantages to a world without belief in moral responsibility. We will have an improved ability to look more clearly at the causes and more deeply into the systems that shape individuals and their behaviour. This would mean getting beyond the visceral desire to blame, punish, and shame, and instead adopt more humane and effective interpersonal attitudes and approaches to criminal justice and social policy.

## 1.4 FREE WILL AND MORAL RESPONSIBILITY

Contemporary theories of free will can be divided into one of two categories, those who defend the reality of free will and moral responsibility and those who are sceptical of it. The first category includes libertarian and compatibilist accounts of free will. Both types of view defend the reality of free will, although they disagree about its nature. The second category includes a group of sceptical views that all take seriously the possibility that human beings do not have free will, and that they are therefore not morally responsible in a way that would make us deserving of blame and praise. The main distinction between the two pro-free will positions, libertarianism and compatibilism, is best understood in terms of the classical problem of free will. The problem of free will comes in trying to reconcile our intuitive sense of free will with the idea that our choices and actions may be causally determined by impersonal forces over which we have no control (Caruso, 2013). For our purposes we can treat determinism as being roughly the thesis that all events and actions, including human action, are the inevitable outcome of prior causes and the laws of nature (Nadelhoffer et al., 2014).

Libertarians and compatibilists respond to the classical problem of free will in different ways. Libertarians believe that free will is incompatible with determinism and that since human beings sometimes exercise free will, determinism must therefore be false. Thus, libertarians reject determinism and defend a contra-causal conception of free will in order to preserve what they believe is a necessary condition for free will – i.e., the ability to do otherwise under exactly the same set of conditions. Compatibilists, on the other hand, defend a less demanding type of free will, one that can be reconciled with determinism. For them, having free will is simply a matter of acting in accordance with our desires and values, free from constraint and compulsion (Caruso, 2013). So even if our behaviour is the inevitable outcome of prior causes, people still deserve praise and blame, reward and punishment, just insofar as they follow their inclinations, are not coerced or restrained, and are in full possession of their cognitive faculties. For compatibilists, if determinism prove true, this would not require us to revise our ordinary

ways of thinking about agency and moral responsibility (Clark, 2007; Monroe, Dillan & Malle, 2014).

Sceptical views stand in contrast to these pro-free will positions by either doubting the plausibility of free will or flat out denying its existence. In the past, the standard argument for free will scepticism was *hard determinism*, the view that determinism is true and incompatible with free will and basic desert moral responsibility. This was either because free will is inconsistent with one's being the "ultimate source" of action (source incompatibilism) or because it prohibits the *ability to do otherwise* (leeway incompatibilism). For hard determinists libertarian free will is an impossibility because human actions are part of a fully deterministic causal web, and as such, "compatibilism is simply operating in bad faith" (Caruso, 2013, p. 23). While hard determinism has few defenders these days – mainly due to accepting interpretations of quantum mechanics that undermine the thesis of universal determinism – more recently a group of contemporary philosophers who refer to themselves as *desert sceptics* have put forward additional arguments against free will and moral responsibility that are agnostic about determinism (Caruso, 2013).

Most sceptics hold that while determinism is incompatible with free will and moral responsibility, so too is indeterminism, especially the variety posited by certain interpretations of quantum mechanics (Pereboom, 2001, 2014). Others argue that regardless of causal laws, we can never be truly free and morally responsible because of the pervasiveness of *luck* (Levy, 2011). Galen Strawson (2010) argues that free will and moral responsibility are incoherent concepts, since to be free in the sense required for moral responsibility we would have to be a *causa sui* (i.e., the cause of oneself) and this is impossible. The thing that all these arguments have in common, and also what they share with classical hard determinism, is the thesis that, "what we do and the way we are is the result of prior circumstances beyond our control. Therefore, we can never be responsible for our actions in the *basic desert sense* – that is the sense that would make us truly deserving of blame and praise." (Caruso, 2018, p. 4).

While there is significant disagreement among philosophers about how the term "free will" should be understood, it is generally agreed that it is closely related to moral responsibility. In fact, one might reasonably think that when philosophers engage in the free will debate, it is really moral responsibility with which they are chiefly concerned

(i.e., whether agents who are causally determined by impersonal forces over which they have no ultimate control can ever be morally responsible for their actions). Galen Strawson (1994, p. 8) writes “It is a matter of historical fact that concern about moral responsibility has been the main motor – indeed the *ratio essendi* – of discussion of free will.” Indeed, the dominant definition of free will in the current literature is one that defines the concept in terms of the control in action needed for moral responsibility. Speaking to this point, Manuel Vargas (2013, p. 2) writes, “I treat free will as the variety of control distinctively required for agents to be morally responsible.” And Eddy Nahmias (2014) insists we should be concerned primarily with free will understood as the set of powers required to be morally responsible, and thereby to be deserving of blame or praise, punishment or reward.

This raises the question of how best to understand moral responsibility within the context of the free will debate. Although there is still considerable disagreement among philosophers on this issue, there appear to be two basic types of moral responsibility that free will is said to ground. These distinct notions of responsibility can be referred to as *consequentialist-based* and *desert-based*.

Consequentialist-based approaches are forward-looking in the sense that agents are considered proper targets of reprobation or punishment for immoral actions on the grounds that such treatment will, say, prevent the agent (or other agents) from performing that type of action in the future. Desert-based responsibility, on the other hand, is considered to be backward-looking and retributivist in the sense that any punitive attitudes or treatments that are deemed appropriate responses for an immoral act/decision are warranted simply by virtue of the action/decision itself, irrespective of whatever good or bad results might follow from the punitive responses (Caruso, 2018b, p. 3).

It is important here to note the critical connection that exists between desert-based moral responsibility and reward and punishment. To say that one is morally responsible for a good or bad act (in the sense relevant to free will) is to say that one *justly deserves* to be rewarded or punished, praised or blamed for that act. (Caruso & Morris, 2016; Pereboom, 2001). Leading retributivist Michael S. Moore echoes these sentiments and highlights the purely backward-looking nature of desert:

[R]etributivism is the view that we ought to punish offenders because, and only because, they deserve to be punished. Punishment is justified, for a retributivist, solely by the fact that those receiving it deserve it. Punishment may deter future crime, incapacitate dangerous persons, educate citizens in the behaviour required for a civilised society, reinforce social cohesion, prevent vigilante behaviour, make victims of crime feel better, or satisfy the vengeful desires who are not themselves crime victims. Yet for a retributivist these are a happy surplus that punishment produces and form no part of what make punishment just: for a retributivist, deserving offenders should be punished even if the punishment produces none of these other, surplus good effect (2010, p. 153).

The “concept of desert to which retributivists appeal is *basic* in the sense that it is not reducible to consequentialist considerations or to forward-looking goods such as the safety of society or the moral improvement of the offender” (Caruso, 2018, p. 2). For the retributivist, because human beings are considered morally responsible in the *basic desert sense*, society is justified in giving them their *just deserts* by way of punishment for moral and legal wrongs. Many prominent philosophers from the various camps in the free will debate have argued that it is the retributivist desert-based moral responsibility that is central to the primary debate regarding free will (Pereboom, 2001; Caruso, 2018; Caruso & Morris, 2016; Strawson, 1994; Waller, 2017).

In support of this view, philosophers frequently claim that retributive attitudes and belief in just deserts are intimately connected to folk conceptions of free will. Nadelhoffer & Tocchetto, (2013, p. 128) write, “it makes a priori sense that people who believe more strongly in free will would be more interested in giving wrongdoers their just deserts.” Smilansky (2000, p. 26) makes the point that “Most people not only believe in actual possibilities and the ability to transcend circumstances, but have distinct and strong beliefs that libertarian free will is a condition for moral responsibility, which is in turn a condition for just reward and punishment.” If these philosophers are correct, it implies that lay people think exercising one’s free will is a necessary condition for one’s being an appropriate target of retributivist treatment. As we shall see, these claims have been lent significant support from recent empirical research.

In building my case against the putative benefits of believing in free will, I will rely primarily on the argument that retributive attitudes, as well as being unnecessary for well-functioning systems of justice, are detrimental to society. For example, there is



substantial evidence to support the view that punishments motivated by factors other than retribution are more effective for attenuating antisocial behaviour (Ward et al., 2012). Take for example, the consequences of a highly retributive criminal justice system like that of the U.S and compare them with those of systems that are more concerned about the welfare and rehabilitation of offenders, such as those found in Scandinavia. The rates of recidivism in those countries (20 – 30%) are approximately half of what they are in the U.S. (40 – 70%) (Ward et al., 2012). Further, while the U.S comprises approximately 4.4% of the world's population it is responsible for a whopping 22% of the world's prison population, giving the U.S. the highest incarceration rate in the world (Walmsley, 2013). Relative to less retributive systems, the high incarceration and recidivism rates found in the U.S. suggest that retributive punishment goals are not necessary for a criminal justice system to effectively employ consequentialist aims of punishment such as rehabilitation, deterrence, and public safety (Morris, 2018).

If it turns out that common sense belief in free will is significantly related to the retributive attitudes that hold in place highly punitive systems of justice, such as that in the U.S, then rather than bringing about negative consequences, leaving free will behind could potentially free us from a number of harmful beliefs, attitudes, and practices (Caruso, 2018a). In light of this, it seems incumbent on those who recommend persevering ordinary belief in free will and desert-based moral responsibility to provide a compelling case for how perpetuating these beliefs would yield better results than exposing the public to arguments undermining them.

## 1.5 SUMMARY

Many researchers have argued that libertarian contra-causal free will is an illusion that is growing less and less believable with increasing evidence supporting mechanistic causes of human behaviour. As these sceptical arguments move beyond academia and into the public domain, questions and concerns about the psychological and societal consequences of reduced belief in free will have emerged. Recent studies in social psychology have indicated that disbelief in free will is related to a range of anti-social behaviours, such as cheating, unjustified aggression, and less helping behaviour. In response to these findings,

*illusionist* philosophers have recommended that even if scientists somehow conclusively showed that free will does not exist it might nevertheless be beneficial to foster widespread belief as a useful fiction. Conversely, because of its close connection with desert-based moral responsibility and retributive attitudes, *disillusionists* maintain that belief in free will has a dark side we would be better off without. And to the extent that lay intuitions and beliefs about the nature of human cognition and moral responsibility are mistaken, philosophers and psychologists ought to do their part to educate the public. With these distinctions in mind, we can turn to consider the moral and psychological implications, as well as the strength of the empirical evidence for both sides of the argument.

## 2. THE EMPIRICAL STUDIES

In this chapter I will show that while belief in free will might be important for sustaining at least some moral attitudes and behaviour, powerful criticisms of the supporting evidence show that it is still not clear to what extent scepticism leads to anti-social behaviour. In light of this, I argue that the illusionist recommendation for “insulating the public” against the putative dangers brought about by exposure to deterministic messages would be too hasty. Moreover, recent evidential support for sceptical arguments claiming that belief in free will is closely connected to belief in just deserts and retributive punishment goals also suggest, as laid out by optimistic sceptics Pereboom and Waller, that leaving free will and desert-based moral responsibility behind may not be as disastrous as many have supposed. There is reason to think that a societal shift away from endorsing free will and moral desert would still leave the functional benefits of punishment intact while avoiding the unnecessary human suffering and economic costs associated with retributivism. So, considered from a consequentialist-based moral perspective, recent empirical findings raise the possibility that scepticism about free will may indeed have positive consequences and even be associated with prosocial attitudes and behaviour.

## 2.1 EVALUATING FREE WILL BELIEFS AND MORAL BEHAVIOUR

The most serious concern about the prospect of widespread acceptance of free will scepticism that illusionist philosophers and psychologists have raised is that empirical research indicates diminished belief in free will encourages anti-social tendencies and decreases prosocial tendencies. This concern has been largely fuelled by two widely cited studies in the social psychology literature – Vohs and Schooler 2008 and Baumeister, Masicampo and De Wall, 2009. The first attempt to study the empirical consequences of weakened belief in free will using experimental methods was undertaken a decade ago by Kathleen Vohs and Jonathan Schooler (2008). In their now classic paper, Vohs and Schooler conducted two experiments in which they manipulated participant belief in free will and then provided them with an opportunity to engage in unethical behaviour. Based on previous findings that show reducing people’s feelings of personal responsibility and sense of control also reduces the amount of effort they exert towards improving their performance on a difficult task.<sup>4</sup> The authors hypothesised that “advocating a deterministic world view that dismisses *individual causation* may similarly promote undesirable behaviour” (2008, p. 49-50). Such expectations appear to be premised on the idea that if people learn that their behaviour is caused by factors outside of their control, they might reason that they can hardly be blamed for their unethical actions. And if they cannot legitimately be blamed for their unethical actions then there is less need for compunction about committing them in the first place.

In one experiment, Vohs and Schooler randomly assigned thirty college students to one of two conditions. Subjects in an anti-free will condition read an excerpt from *The Astonishing Hypothesis*, a book by noble-prize winning scientist Francis Crick (1995) claiming that rational high-minded people, including most scientists now recognise free will is an illusion. The control group also read a passage penned by Crick but one that was unrelated to free will. Using an unpublished scale by Paulhus and Margesson (1994),

---

<sup>4</sup> Vohs and Schooler cited previous studies that indicate a sense of personal responsibility motivates people to modify their behaviour to align with their beliefs. For example, in one study participants who were told outcomes are the result of concerted effort, rather than natural intelligence, were more likely to persevere with solving a difficult problem (Mueller et al, 1998).

the researchers confirmed that participant free will beliefs in the experimental condition had shifted in the predicted direction. The participants were subsequently asked to solve math problems on a computer. However, they were told that owing to a technical glitch, the answers would pop-up on the screen after the problem if they did not hit the space bar. They were asked not to cheat but told no one would know either way. Consistent with their hypothesis Vohs and Schooler reported significantly more cheating in the anti-free will condition compared to controls. They concluded that if exposure to deterministic messages has the effect of encouraging people to behave unethically, then finding ways to insulate the public against such information is “imperative” (2008, p. 54).

In the second influential study, Roy Baumeister and colleagues (2009) sought to extend the Vohs and Schooler findings into a broader context, namely helping and aggression. The researchers predicted that participants with weakened free will belief would be less helpful and show increased aggression toward someone they were induced to dislike. They reasoned that belief in free will is crucial for the sense of responsibility required to motivate people to control their otherwise selfish impulses in favour of more prosocial behaviour. They believed exposure to deterministic statements claiming *all* human behaviour is determined by impersonal forces would give people the impression that their choices and actions are *causally inefficacious* and it would be futile exerting mental energy trying to restrain such impulses. On this reasoning, Baumeister hypothesised that “disbelief in free will reduces prosocial tendencies” (2009, p. 261).

Undergraduate students read through fifteen sentences either asserting human freedom of action or deterministic inevitability. Helping behaviour was measured as willingness to help others in a number of hypothetical scenarios presented in a series of vignettes (e.g., giving money to homeless person, allowing fellow classmate to use one’s cellular phone), and aggression was measured as the amount of hot-sauce participants served to a person they were led to believe had wronged them and who had said explicitly they didn’t like it. Results supported Baumeister’s hypothesis. Inducing people to disbelieve in free will led to an increase in aggression and a reduction in willingness to help. Baumeister concluded that, “Volition and self-control require the person to expend energy, and these expenditures enable them to act pro-socially. Apparently, disbelief in free will subtly reduce people’s willingness to expend that energy. Hence, disbelief in

free will serves as a cue to act on impulse, a style of response that promotes selfish and impulsive action such as aggressing and refusing to help” (2009, p. 267).

More recently Baumeister was part of a research group (Crescioni et al., 2016) that found believing in free will is associated with a large number of traits and beliefs commonly studied in positive psychology. The researchers conducted a correlational study using survey data to identify what other beliefs and traits are linked to variations in laypersons belief about free will. In particular, they tested hypotheses linking beliefs about free will to traits and beliefs involving agency, happiness, meaningfulness, and general positivity. Results indicated that individuals with stronger belief in free will were happier and more satisfied with life, found life to be more meaningful, were more mindful, and expressed more gratitude toward others. The researcher concluded that, “Insofar as future research may begin to change the collective understanding of free will, the implications may ripple through society and alter human happiness, meaningfulness, agency, and other factors. In addition to being a profoundly important metaphysical question, whether one believes in free will appears to be an important social reality” (2016, p. 61).

As a first observation, it is important to notice that the experimental studies of both Vohs and Schooler, and Baumeister and colleagues did not only suggest a link between belief in free will and less prosocial/greater antisocial behaviour. They also suggested a link between a *deterministic-world-view* and such tendencies. For example, on the basis of their studies in which Vohs and schooler manipulated belief in free will with statements describing a deterministic worldview, they conclude that “widespread encouragement of a deterministic worldview may have the inadvertent consequence of encouraging cheating behaviour” (Vohs & Schooler, 2008, p. 53). Both groups of researchers expected statements describing deterministic ideas to induce disbelief in free will. From the fact that these studies suggest a three-way relationship between a deterministic world view, weaker belief in free will, and an increase in unethical behaviour Stephen Morris (2018) notes some critical points.

To begin with, given the sort of the statements that these researchers used to manipulate belief in free will, it seems clear that they were aiming to induce disbelief in libertarian contra-causal agency. Recall, libertarians believe that free will is incompatible with determinism and that since human beings sometimes exercise free will, causal

determinism must be false. The following two statements that both groups employed suggest that they expected an increase in unethical behaviour to be elicited by way of the subjects taking them to imply the non-existence of libertarian free will “A belief in free will contradicts the known fact that the universe is governed by lawful principals of science” (Vohs & Schooler, 2008, p. 51); “All behaviour is determined by brain activity, which in turn is determined by a combination of environmental and genetic factors” (Baumeister, 2009, p. 262). Moreover, given the description of free will endorsed by Baumeister elsewhere in the literature – e.g., “a pervasive belief in an elusive, metaphysically and scientifically mysterious phenomenon” (Clark, Baumeister & Ditto, 2016, p. 210) – it is hard to doubt that belief in libertarian contra-causal free will is the intended target for manipulation.

Given this focus, the Vohs and Schooler, and Baumeister findings indicate that information calling the contra-causal assumption into question has adverse behavioural consequences. What can be said about Vohs and Schoolers conclusion that it’s “imperative” that we insulate the public against the dangers of unethical behaviour generated by deterministic messages? On balance, it is difficult to imagine what kind of reasonable measures could be taken to avoid this threat. As Morris (2018) notes, if we are to take the admonitions of Vohs and Schooler seriously, we would have to consider whether or not to stop science advancing. Or at least cease communicating findings from neuroscience, psychology, and other fields of inquiry that continue to make discoveries that show libertarian free will to be increasingly dubious. Moreover, assuming that cutting off communication between scientific information and the public were even feasible, it should take a lot of supporting evidence demonstrating the detrimental effects of undermining belief in free will to justify such. In the following I show that the findings of Vohs and Schooler, and Baumeister and colleagues fall far short of the epistemic standard of proof that should be required for such a drastic move.

## 2.2 DETERMINISM, FATALISM, AND THE PROBLEM OF PRIMING

While the findings discussed in the previous section appear to support illusionist concerns over the anti-social consequences of dispensing with belief in free will, caution must be advised in drawing any firm conclusions. Critics have noted significant flaws in

the methodology of the studies calling into question any long-term potential for unethical behaviour due to exposure to deterministic messages (Pereboom, 2014; Slattery, 2014, Nadelhoffer & Wright, 2018). First of all, the passages used to prime disbelief in free will appear to be priming the wrong thing. It has been noted that rather than priming free will scepticism, the Crick excerpt is actually priming a scientific reductionist view of the mind, one that is said to demonstrate that free will is an illusion (Pereboom, 2014; Pereboom & Caruso, 2018). For example:<sup>5</sup>

You, your joys and your sorrows, your memories and your ambitions, your sense of personal identity and free will, are in fact no more than the behavior of a vast assembly of nerve cells and their associated molecules. Who you are is nothing but a pack of neurons ... although we appear to have free will, in fact, our choices have already been predetermined for us and we cannot change that.

The Crick prime may be giving subjects the mistaken impression that scientists have concluded that their beliefs, desires, and choices are causally inefficacious – a claim not held by most philosophical sceptics.<sup>6</sup> Free will sceptics do not deny that we are *causal agents* who make decisions, and engage in deliberative processes and reasoning. Instead, they claim that these acts themselves are the result of causal factors that are beyond the agent's control (Pereboom, 2001). It is important then, that the researchers should prime the correct belief and not the impression that scientists have dismissed the idea that human action has causal power. Both Vohs and Schooler, and Baumeister and colleagues expected that deterministic messages would give subjects the impression that their choices and actions are causally inefficacious. That is, they expected the primes to induce a sense of fatalism. The fatalist believes that it is futile to act upon higher level motives, since the future is already fixed. Vohs, Schooler and Baumeister were exploiting the possibility that lay people would associate determinism with fatalism and this would in turn encourage the “why bother mentality” that would produce anti-social behaviour (Vohs & Schooler, 2008, p. 54). As Tom Clark (2013) has noted, if people come to believe they do not have ultimate control, and if they have something like the authors

---

<sup>5</sup> This excerpt of the Crick prime was taken from the supplementary material provided by Shariff et al (2014).

<sup>6</sup> This criticism has been made by Eddy Nahmias on the Garden of Forking Paths Blog ([http://gfp.typepad.com/the\\_garden\\_of\\_forking\\_pat/2008/01/on-the-benefits.html](http://gfp.typepad.com/the_garden_of_forking_pat/2008/01/on-the-benefits.html)).

(mis)conception of what not having it entails, then indeed they might become demoralized. This could explain the results of the study.

If subjects interpreted the Crick passage and other deterministic statements as entailing fatalism, then since this goes beyond what is entailed by determinism, the negative behaviours might not be due to increasing belief in the latter. As such, the results of these studies do not necessarily indicate that propagating a deterministic worldview or a sceptical attitude will have an adverse effect on people's behaviour. So long as endorsements of a deterministic worldview and free will scepticism go along with the notion that neither view entails the causal inefficacy of acting in accordance with our beliefs and desires then any negative impact from exposure to determinism and sceptical arguments might be avoided (Morris, 2018). As Eddy Nahmias puts the point, "no one has shown that telling people they lack just what philosophical (not scientific!) Sceptics say they lack, and nothing more has any bad effects on behavior."<sup>7</sup>

Relatedly, over four studies Nadelhoffer and Wright (2018) employed the same priming techniques as Vohs and Schooler and Baumeister and colleagues to manipulate belief in free will. After their initial primes failed to have any effect, they decided to try the Crick prime, which in their words "As far as anti-free will primes go this is the bludgeon" (2018, p. 288) when it comes to anti-free will primes. Only with this prime did people show a decrease in free will belief. However, Nadelhoffer and Wright (2018) make the point that the problem with this prime is that "it not only challenges free will, but it also challenges dualism, the soul, self-awareness, and even choice itself" (2018, p. 288). Similarly, Eddy Nahmias, in describing his difficulty replicating priming effects writes: "the effects don't always replicate, and they only seem to work with over-the-top primes that suggest all kinds of threats to agency."<sup>8</sup> So, although some studies have had success using primes in free will research (Crescioni et al., 2016; Clark et al., 2017) the results have been mixed and hard to replicate. Furthermore, contrary to what one might conclude in light of published work in social psychology, there is reason to expect that people's deep seated intuitions about something as fundamental as having free will would be fairly resistant to simple manipulations (Nadelhoffer & Wright, 2018).

---

<sup>7</sup> Eddy Nahmias made these comments on the blog Flickers of Freedom on 3/18/2015: [http://philosophycommons.typepad.com/flickers\\_of\\_freedom/2015/03/free-will-skepticism-just-world-belief-andpunitiveness/comments/page/1/#comments](http://philosophycommons.typepad.com/flickers_of_freedom/2015/03/free-will-skepticism-just-world-belief-andpunitiveness/comments/page/1/#comments)

<sup>8</sup> Ibid



To make this point salient, consider an analogy suggested by Trick Slattery (2014). Suppose I were to design a study that compared how atheists and theists responded to action X. Further, suppose I do this not by assessing whether my subjects are *actually* atheists or theists, but instead, I take thirty random people (who statistically will almost all be theists of some kind) and have half of them read a text claiming God does not exist while the other half read something else. Imagine that this elicits a slightly more atheistic response from the subjects who read the anti-God statements. Not only would most or all of these people not be atheists if asked directly, but even if the anti-God statements produced a little hesitation and dissonance with regard to their theism, it would still be response of a theist to an atheistic message. The fact that most theists have some sort of negative opinion or bias regarding atheism would most likely play into the psychology. So, when we test for how they respond to action X, what we elicit is not generally the response of an atheist but the response of an (at best) temporarily confused theist with all their biases still intact. This is analogous to people who believe that they have free will: they will be likely to hold all manner of incorrect biases related to this belief, even if they are suddenly told that they lack free will.

With that in mind, we should infer that the test subjects did not believe there was *no* free will, only that their pro-free will beliefs were slightly weaker than the controls. It is, therefore, misleading to frame this in terms of inducing *disbelief* in free will. Nadelhoffer and Wright (2018) note that it is not clear that any study that has been conducted thus far has been able to move people's free will beliefs from above the midline to below the midline. And this suggests that while challenging people's beliefs in free will may weaken belief, it does not undermine such beliefs. However, you would not think this if you looked at how the results in the extant literature have been presented and discussed.<sup>9</sup>

---

<sup>9</sup> Nadelhoffer and Wright (2018, p. 272) concluded that, "In this respect, our work should serve as a cautionary tale for philosophers, psychologists, and pundits who want to discuss the potential ramifications of the supposed death of free will. For while it's certainly possible for people to change their mind about free will, it's not clear that researchers have figured out effective methods for bringing these epistemic changes about (even temporarily)."

## 2.3 FAILURE TO REPLICATE AND UNVALIDATED MEASUREMENT TOOLS

Leaving the problems of priming disbelief in free will behind for the moment, let us take a look at the problems of non-replicability, small sample sizes, and unreliable measurement tools. Subsequent studies have had a difficult time replicating the Vohs and Schooler findings, and the study was also involved in the so-called replication crisis in social psychology (Carey & Rolston 2015; Zwaan, 2013). *The New York Times*, for example, published an article focusing on the failure to replicate these findings since it was the most cited out of the hundred studies in the Reproducibility Project (Carey & Rolston 2015). However, even before the Reproducibility Project failed to replicate the findings, a large study undertaken at the university of Rotterdam had already failed to replicate between group differences in strength of free will belief and effect on cheating (Zwaan, 2013). One possible explanation is that the Vohs and Schooler studies were underpowered. The researchers relied on small convenience samples (e.g., 30 students) whereas Zwaan, used 150 subjects. Because of the high variability between subjects in small samples the effect sizes are often large and do not reflect a realistic between group difference. This means that when studies are underpowered, the results are likely to be inconsistent and misleading, and generalisability cannot be assumed.

Finally, it is difficult to overstate the importance of using valid and reliable measurement tools in empirical research. Unfortunately, the measurement tool used to examine belief in free will and determinism in all the aforementioned studies, *The Free Will and Determinism Scale* (FAD-4) was an unpublished scale by Paulhus and Margesson (1994) with known methodological weaknesses (as discussed in chapter 3).

In sum, the effects of priming, failure to replicate, insufficient sample sizes, and use of an unreliable measurement tool shows the validity and reliability of the Vohs and Schooler, and Baumeister et al. findings to be highly questionable. Furthermore, even if it were the case that exposure to deterministic messages elicited serious anti-social behaviours, this cost would have to be compared to the benefits one might expect from any upsides to accepting a sceptical world view. Indeed – and critically for the empirical work I pursue here – there is a small but growing body of evidence indicating that certain *positive* effects may follow from accepting a sceptical world view. If these effects are shown to be robust and replicable, then they too need to be considered before drawing any firm conclusions about the overall value of believing in free will. To shed more

empirical light on the illusionism debate my primary goal in the next section is to survey the recent findings from social psychology and experimental philosophy to show the potential dark side of believing in free will.

## 2.4 AN UPSIDE TO DESERT SCEPTICISM

One reason to think that agents might not deserve to suffer for moral wrongs they have committed in the basic desert sense is that they are not free and morally responsible in the sense required. As Caruso (2018, p. 2) puts the point “Free will scepticism maintains that what we do and the way we are is ultimately the result of factors beyond our control and because of this we are never morally responsible for our actions in the desert based sense – the sense that would make us truly deserving of blame and praise.” Desert sceptics are motivated by the intuition that it’s inappropriate to inflict certain types of rewards or punishments upon agents who are causally determined by forces over which they have no ultimate control. On this reasoning, unlike free will believers who tend to endorse more severe retributively motivated punishment (Carey & Paulhus, 2013) desert sceptics believe punishment should be based on forward-looking consequentialist considerations such as rehabilitation, deterrence, and wider goods such as the safety of society.<sup>10</sup>

Sceptics maintain that the driver of retributive attitudes is the belief that individuals *justly deserve* what they get. Moreover, sceptics claim that the idea of just deserts that’s central to desert-based moral responsibility is a pernicious one as it often encourages punitive excess in criminal justice and perpetrates social and economic inequalities (Waller, 2011). On the ordinary interpersonal level desert-based moral responsibility justifies reactive judgments such as indignation, resentment, and a certain type of moral anger that is corrosive to our personal relationships (Pereboom, 2001).

---

<sup>10</sup>Here Caruso reiterates the retributive motivation and speaks to the unjustness of basic desert: “The retributivist justification for punishment maintains that punishment of a wrongdoer is justified for the reason that she deserves something bad to happen to her because she has knowingly done wrong. For the retributivist, it is the *basic desert* attached to the criminal’s immoral action alone that provides the justification for punishment. This is to insist that to hold people truly or ultimately morally responsible for their actions – that is, to hold them responsible for the results of the morally arbitrary in a non-consequentialist desert-based sense – for what is ultimately beyond their control, which is fundamentally unfair and unjust” (2018, p. 4).

Therefore, belief in free will, rather than providing the pragmatic benefits as many claim, is too often taken to justify treating people in severe and demeaning ways. For these reasons Waller, Pereboom, Caruso and other desert sceptics argue that belief in free will and desert-based moral responsibility, rather than being a good thing, actually has a dark side we would do well to leave behind.

Recent research that examines lay belief about free will lends strong support to the sceptical claim that such beliefs contain a strong retributive element. Social psychologists Jasmine Carey and Delroy Paulhus (2013) conducted a correlational study using a more recent improved version of the Free Will and Determinism Scale (FAD-Plus; Paulhus & Carey, 2011) – a questionnaire used to measure people’s beliefs and attitudes about free will and related concepts – along with measures of religiosity, political conservatism, just world beliefs, and right wing authoritarianism.<sup>11</sup> Based on traditional assumptions about conservatism and punishment (e.g., Adorno et al., 1950., Altemeyer, 1973) and previous research suggesting conservatives demonstrate increased punitive and blaming behaviour, Carey and Paulhus (2013) hypothesised that the personal responsibility ethic and just deserts ideology emphasised by conservatives necessitates a firm belief in free will.

Their predictions were largely borne out by the data. Stronger belief in free will was associated with higher levels of religiosity, a conservative worldview, right wing authoritarianism, and just world belief. Moreover, consistent with the responsibility theme was the finding that free will believers endorse harsher criminal punishments motivated by retribution rather than consequentialist goals of rehabilitation, deterrence, and public safety. Carey and Paulhus concluded that “free will belief relates to the expectation that people will control their own impulses and criticize others for not doing the same. If others misbehave, those high in free will belief are willing to apply the necessary sanctions that is punishment” (2013, p. 139).

---

<sup>11</sup> The RWA scale was developed to measure the degree of obedience to authority and adherence to existing social conventions. As a behavioural trait right wing authoritarianism is typically defined as submission to established and legitimate authority, sanctioned general hostility towards various persons, and adherence to majority endorsed social conventions. Scores on RWA have been found to reliably correlate with attitudes of punitiveness, racial prejudice, homophobia, religious orthodoxy, and victim blaming.

Nadelhoffer and Tochetto (2013) sought to replicate the findings using a slightly different scale – the Free Will Inventory (FWI; Nadelhoffer et al., 2014). Based on Paulhus and Carey’s findings on worldview implications of believing in free will and research suggesting different psychological motivations and tendencies underlie political orientation (Jost, 2006), they predicted higher scores on the (FWI) free will subscale would be associated with higher scores on right wing authoritarianism, just world beliefs, and several other measures of beliefs and attitudes about social justice.<sup>12</sup> Results supported Carey and Paulhus’s (2013) finding that belief in free will is associated with right wing authoritarianism and just world belief. Given the consistency of findings between these independent research groups, it’s important to highlight just how troubling some of these correlations between believing in free will and political ideology happen to be (Caruso, 2018). Consider, for instance, a few sample items from the Just World Belief Scale (JWB; Lerner, 1980) that Carey and Paulhus used to validate the FAD-Plus:

- By and large, people deserve what they get.
- I firmly believe that injustices in all areas of life are the exception rather than the rule.
- People who meet with misfortune have often brought it on themselves.

For many people, these items will express concerning and perhaps potentially dangerous ideas. Yet based on the findings of Carey and Paulhus (2013) and Nadelhoffer and Tocchetto (2013), it appears these sorts of beliefs are closely related with belief in free will. To make clear the potential dark side of belief in free will and moral responsibility, Nadelhoffer and Tocchetto describe the origins and associations of the JWB:

In order to measure the degree to which persons are willing to believe that everyone deserves what happens to them, Lerner (1980) developed the JWB scale. Scores on the scale have been found to correlate with the presence of frail religious beliefs, internal locus of control, and with the likelihood of derogating innocent victims. In addition, people who score high on JWB are more likely to trust current institutions and authorities, and to blame the poor and praise the rich for their respective fates (2013, p. 132).

---

<sup>12</sup> The study also employed the Social Dominance Orientation Scale (SDO; Pratto, Sidanius, Stallworth, & Malle, 1994), and measures of religiosity, and political conservatism (not cited).

Caruso (2018a) frames belief in a just world as a blame-the-victim-approach that grounds the idea that people deserve what they get and those who meet with misfortune have mostly brought it on themselves. Empirical evidence suggests that although people generally have a difficult time accepting injustice, this is particularly true of those with a stronger belief in a just world. Appelbaum (2002) has demonstrated that when there is limited action that can be taken to alleviate injustice those with stronger just world belief are more likely to derogate victims, so as to reconstruct the situation to make it appear just. By doing this, the belief that the world is just can be preserved and the perception of order and stability maintained.

The idea is that belief in a just world allows for the perception of immunity from negative consequences if one refrains from doing anything to deserve such consequences. Given the psychological benefits, belief in a just world continues to exercise powerful, often unconscious, effects on our attitudes about free will and moral responsibility (Caruso, 2018a). Yet, any existential benefit this belief affords us appears to be cashed out at a high price. Waller (2015) notes, ironically, the costs of belief in a just world are paid in fundamental *injustice*. This claim is supported by empirical evidence showing that people with stronger belief in a just world fail to recognise the full extent of social inequalities and are less likely to view extant inequality as unfair or unjust compared to those with weaker just world belief (Smith, 1985). We can see evidence of just world belief in the regrettable but all too common tendency to blame rape victims for their circumstances. When there is no way, or it is too costly, to help innocent victims our belief in a just world is severely threatened and the most convenient way of defending that belief is to change the status of the victim from innocent to guilty (Caruso, 2018a).<sup>13</sup>

---

<sup>13</sup> Waller (2013, p. 73) articulates this unfortunate tendency: “The case of rape victims is the most obvious and extensively studied example of this phenomenon. Rape is a brutal, demeaning, and trauma-producing crime; in a just world, no innocent person would be subjected to such a horrific fate. Thus, there is a powerful tendency to see rape victims as really not quite so innocent: they dress provocatively; they were “loose” women; they did something to put themselves in that situation (they were careless about where they walked, or they drank too much); they “led him on” or were “asking for it” (thus in some parts of the world, rape victims are subject to death by stoning). Harsh cross-examination of those who claim to be rape victims are notoriously common; those harsh cross-examinations are common because they are often effective; and they are often effective because juries—eager to preserve their belief in a just world—are already inclined to see the victim of this terrible ordeal as other than innocent.”

## 2.5 TWO COMMENTARIES ON THE SOCIAL IMPLICATIONS OF BELIEF IN FREE WILL

Waller's explication of just world belief and its relation to attitudes about rape victims is just one of the many unfortunate examples of the pernicious nature of belief in a just world. Other examples include blaming those in poverty for their own circumstances, maintaining that offenders deserve harsh punishments, assuming beneficiaries are simply lazy, and blaming parents and children for poor educational outcomes (Caruso, 2018a). So, what are the corollary social and moral implications of believing in free will?

In June 2009, the Joseph Rowntree Foundation published findings showing that 69% of UK citizens agreed that virtually everyone remains in poverty in Britain not as the result of social disadvantage or biological disability but through *choice*. Those studied tended to attribute success or failure overwhelmingly to the individual rather than structural factors. Further, because of their belief in the basic fairness of deserved inequalities the respondents were found to be almost completely unconcerned with the idea of supporting greater equality while on the other hand asserting that Britain is a beacon of fairness with opportunities for all. The Rowntree researchers reported that there was a clear sense across all of the groups surveyed that an individual's situation is largely of his or her own making (Bamfield & Horton, 2009).

James B. Miles (2015) argues that the Rowntree Report findings may be taken as evidence that in Britain, and very likely in America,<sup>14</sup> 69% to 83% of the population are ultimately calling on free will to excuse indifference to the poor. According to Miles, "free will may just be the primary excuse many use to legitimise a contempt for the poor that would exist independent of their professed belief in free will, but free will assertion nonetheless provides the ethical fig leaf for such contempt that would be far harder to rationalise (and therefore tolerate) without the myth of free will" (2013, p. 212). Further,

---

<sup>14</sup> The influential American Pew Research Centre for People and the Press has concluded that successive Pew Global Attitudes polls find that at every level, Americans are far more likely than Europeans to believe that individuals, not society, are responsible for their own failures, economic and otherwise. Allen & Dimock, 2007).

Miles believes that belief in free will does not just excuse indifference to poverty but creates and maintains much of that poverty in the first place.

Likewise, Richard Oerton (2018) offers a blistering criticism of the 2012 American presidential candidate Herman Cain who said: “If you don’t have a job and you’re not rich, blame yourself” (2016, p. 40). He writes that Cain’s statement is ludicrous even on a superficial level. For everyone to be rich is a logical impossibility as well as a practical one. At deeper level it is also vicious as well as ludicrous being fuelled by the “idiotic” idea of unmitigated free choice. Oerton writes that Cain seems to actually believe that all of us can cast off our mental endowments however disabling, our social status however low, our upbringings however abusive, our educations however inadequate, and all other things that have built us to be the way we are, and we only have ourselves to blame if we do not. He thinks, it is a “sad fact that the idea of free will is constantly called in aid, explicitly or implicitly, as justification for the social status quo however unfair and unjust this may be” (2018, p. 40). This leaves us with a social climate whereby sympathy and understanding for the poor, deprived, and otherwise disadvantaged is thereby destroyed or much reduced (along with the obligation of others to do something about their plight) by the idea that they could have freely chosen to escape from their situation and have freely chosen not to. And to add insult to injury, those who’s luck has left them ill-equipped for life are cursed not only by their disadvantage but with the fact that society looks down on them and even blames them for it.

To illustrate common contempt for the poor Oerton draws attention to how the media treats political support for social welfare initiatives. Towards the end of 2015 in the UK, leader of the Labour party Jeremy Corbyn expressed a strong commitment to progressive social policy and was condemned by the media as the party for “down and out losers” the party of the “undeserving” (2018, p. 45). In desert sceptic style Oerton delivers the clear message “that what, in actual fact, the undeserving don’t deserve is the biological and environmental luck that made them the way they are and denied them the motivation and ability to change their lives – that and the condemnation they receive as a result” (2018, p. 45).



## 2.6 A MECHANISTIC VIEW OF HUMAN NATURE REDUCES RETRIBUTION

If Miles and Oerton are right in thinking common sense views about free will are tightly connected to social attitudes that blame the poor for their circumstances and maintain social and economic inequality. Then the claim that society would be better off without the belief in free will should be further bolstered by empirical evidence indicating its close connection with retribution. As we shall see, an examination of lay attitudes about free will lends support to arguments claiming that such beliefs entail a strong retributive element. For instance, in a recent study, Shariff et al. (2014) found that people with relatively weaker belief in free will demonstrated less retributive attitudes regarding punishment of criminals. They also found that learning about the (mechanistically implied) neural basis of human behaviour, whether through experimental manipulations or simply by completing an undergraduate neuroscience course, reduced attributions of blameworthiness and retributive punishment goals in a hypothetical scenario. Shariff and colleagues hypothesised that if free will beliefs are closely linked to attributions of moral responsibility, then reducing these beliefs should make people less retributive in their attitudes about punishment.

To test their hypothesis, Shariff et al. (2014) conducted a correlational study and three experiments aimed at discovering the relationship between free will belief and retributive attitudes. In the first study, using scores from the FAD-Plus and various other questionnaire data, they examined associations between belief in free will and judgments about the relative importance of retribution and consequentialism in motivating punishment goals for criminal offenders. They found that while stronger belief in free will predicted retributive punishment, it was not predictive of consequentialist punishment. The finding that belief in free will is unrelated to the endorsement of consequentialist treatments, though closely connected to retributive treatments, confirms the hypothesis that desert-based responsibility is the operative type of responsibility at work in the lay concept of free will. Results also point to the possibility that “motivation to punish in order to benefit society (consequentialist punishment) may remain intact, even when the need for blame and desire for retribution are forgone” (Shariff et al, 2014, p. 1569). The researchers describe the potential benefits of these findings as follows:

[A] societal shift away from endorsing free will could occur without disrupting the functional role of punishment. Society could fulfil its practical need for law and order, leaving the social benefits of punishment intact while avoiding the unnecessary human suffering and economic costs of punishment often associated with retributivism (2014, p. 1569).

This finding suggests that there is no reason to believe moral chaos would ensue if we gave up retributive justice. To that end, results of this study indicate that if there were to be a reduction in free will belief other justifications for punishment remain intact and unaffected.

In the second study, participants in an anti-free will condition read a passage that explicitly rejected free will. Those in the test condition and controls were then instructed to recommend a prison sentence for a hypothetical man who beat another man to death; they also rated the offender's blameworthiness. The participants in the test condition saw the offender as less blameworthy and recommended roughly half the prison time of controls. This suggests exposure to anti-free will messages elicits a significant reduction in the attribution of moral responsibility and retributive behaviour. The third study aimed to bolster this finding using a much subtler manipulation. The test group simply read a neuroscience article depicting a mechanistic view of human behaviour. Consistent with the second study participants who were exposed to the prime demonstrated reduced belief in free will, blamed the offender less, and recommended roughly half the prison time compared to controls. The fourth study tested students before and after they took an undergrad neuroscience course. Once again results showed a reduction in retributivism after exposure to a mechanistic view of behaviour. Summing up their findings, the researchers concluded that: "Convergent results across a correlational study and three experiments with diverse manipulations consistently demonstrated that shifting from a belief in free will toward a mechanistic view of human behaviour reduces support for retributive punishment" (Shariff et al., 2014, p. 1568).

Besides showing that lay people generally tend to see free will as grounding desert-based moral responsibility and not consequentialist-based notions, the findings of Shariff et al (2014) draw attention to another issue that has been a long-running source of contention between the different camps in the free will debate – namely whether people are natural compatibilists or incompatibilists. If merely exposing subjects to neuroscience articles indicating a mechanistic view nature of human behaviour is enough to

significantly reduce people's belief in free will, it is a strong indication that there is a decidedly incompatibilist element to lay beliefs.

## 2.7 ARE PEOPLE NATURAL COMPATIBILISTS OR INCOMPATIBILISTS?

The long-running dispute between compatibilists and incompatibilists about whether free will and moral responsibility are compatible with determinism – the compatibility question – is one of the most enduring and heated deadlocks in Western philosophy. Recall, (Hard) Incompatibilists hold that people are not morally responsible in the desert-based sense if determinism is true (i.e., if every event is an inevitable outcome of prior causes and the natural laws). By contrast, compatibilists maintain that desert-based moral responsibility is not undermined by the truth of determinism (or indeterminism) and therefore desert based moral responsibility is justified. The Philosophical arguments supporting these two positions have employed a range of different resources, including quantum mechanics, neuroscience, social psychology, and basic metaphysics. However, over the last decade much of the discussion has centred around empirical findings from experimental philosophy that draw on lay intuitions about the compatibility issue. Rather than directly addressing the metaphysical questions, experimental philosophers focus on the natural compatibility question which is concerned with whether people start off as compatibilists or incompatibilists prior to considering the philosophical theories. It has now been more than a decade since experimental philosophers first began using empirical methods to investigate lay intuitions about the compatibility problem. While these studies have made numerous important contributions to our understanding of free will and moral responsibility, the issue of natural compatibility remains unresolved with evidence supporting both sides

The question of natural compatibility first took root in experimental philosophy because of claims made by both compatibilist and incompatibilist philosophers to the effect that, since lay-people are (in)compatibilists, the burden of proof lay on the opposing camp to demonstrate that theirs was the natural position. If it turned out that a significant majority of people make judgments that support one view over the other this would at least give “squatters rights” to whichever position had the empirical support (Nahmias et al., 2005, p. 564). This challenge was also fuelled by a strong sense that the justification

for how we *ought* to think about free will and moral responsibility lay with whichever turned out to be the common-sense view (Nahmias et al., 2005). If it should turn out that lay-people are natural compatibilists Nahmias and colleagues argued (against desert scepticism), an acceptable theory of free will should be able to “elucidate the abilities presupposed by our practices of attributing praise and blame, our expressions of reactive attitudes such as indignation and gratitude, and our systems of punishment and reward” (Nahmias et al., 2006, p. 30). So according to Nahmias (2014), should compatibilism prevail, we should be concerned primarily with free will understood as the set of powers or capacities required to be morally responsible in exactly the desert-based sense.

Experiments were conducted in experimental philosophy similarly for both sides of the argument. Participants were situated within the philosophical debate, usually by reading a vignette, and then instructed to take sides by choosing one of two response options (e.g., whether an agent in a deterministic universe is morally responsible for a transgression or a good deed). Should a participant assent to a statement about moral responsibility that is logically incompatible with determinism, it was concluded that for all intents and purposes she is a natural incompatibilist irrespective of whether she is explicitly aware of incompatibilism. Conversely, if it is judged that an agent is morally responsible in a deterministic scenario then this was interpreted as compatibilism.

However, given the messy results of these dichotomous set-ups, each of the two camps proposed error theories for their opponent’s conclusions. In numerous papers Nahmias and collaborators attempt to explain away incompatibilist intuitions by suggesting that lay people misinterpret determinism as either bypassing or fatalism (e.g., Murray & Nahmias, 2014; Nahmias, Shepard & Reuter, 2014). On the other side, Nichols and Knobe (2007) offer an error theory for compatibilist intuitions arguing that these judgments occur only when affective responses cloud deliberative judgements. To date, there is still considerable disagreement about how best to interpret results from the compatibility experiments. Given the mixed results, Nadelhoffer and Tochetto (2013) concede that for every piece of evidence that seems to support compatibilism as the folk view, there is competing evidence that suggests people find incompatibilism every bit as intuitive.

Nadelhoffer et al (2014) sought more clarity on the (in)compatibility issue and turned to scale methodology commonly used in social psychology hoping the results

would allow for a more fine-grained analysis about the content of lay beliefs and related constructs. The strength of scale methodology is that it is able to distinguish between competing views even when participants agree with both, and even to a similar degree. So, in theory, if it was discovered that stronger belief in free will is associated with weaker belief in determinism (and vice versa), then this is evidence for lay incompatibilism. Two important findings from these psychometric studies have emerged. First, scores on both the FAD-Plus and the FWI have shown belief in free will and belief in determinism are *unrelated* rather than negatively correlated as one would in theory have predicted. This suggests that the two beliefs are driven by largely independent cognitive mechanisms where belief in one is not affected by belief in the other. Second, in the process of attempting to distinguish compatibilist beliefs from incompatibilist beliefs, Nadelhoffer et al (2014) found preliminary evidence that not only do *most* people believe in libertarian contra-causal free will, but stronger belief in such is associated with weaker belief in determinism and vice versa. The latter result was a breakthrough finding. For the first time in psychometric research on free will there was evidence supporting lay incompatibilism. Nadelhoffer and Tochetto (2013, p. 130) had this to say about the finding:

... results reveal that many people explicitly endorse precisely the kinds of libertarian views about free will that are often attributed to the folk by the scientific sceptics and illusionists. So, as we saw earlier, while there is clearly a compatibilist element to how people ordinarily think about free will, it is becoming increasingly clear that there is also an *incompatibilist* element to people's free will belief that is every bit as deeply ingrained ... Our suspicion is that the debate about folk incompatibilism is ultimately based on a false dichotomy – namely that people must either be compatibilists or incompatibilists. It is becoming increasingly clear that people are *both* ...

## 2.8 SUMMARY

In this chapter, I have identified and explicated two gaps in the extant literature on free will. First, the original studies purporting to show the negative consequences of free will scepticism fail to provide conclusive evidence about whether the increase in antisocial behaviour resulting from exposure to deterministic messages is likely to be serious or to last for any significant amount of time. But more importantly, free will belief

researchers did not prime free will scepticism. Rather they expected that participants would associate statements describing a deterministic world-view with a particular sort of fatalism that renders desires and intentions causally impotent. So even if we accept that exposure to deterministic ideas might cause a temporary confusion about what it means to lack free will, it says nothing about the long term effects of a sceptical attitude. As such, the extant findings do not present a compelling case for thinking that relatively low-stakes findings obtained in highly-controlled laboratory conditions generalise to higher-stakes immoral behaviour in the real world. Furthermore, there is strong evidence indicating that belief in free will is associated with destructive tendencies such as blaming victims of misfortune for their own circumstances, intolerance of unconventional life-ways, and retributive attitudes. So if it were true that exposure to sceptical arguments produced some serious unwanted behaviours, this cost would have to be compared to the benefits that could be expected as a result of wide acceptance of free will scepticism. Even if people initially leaned toward irrational ideas that are a non-sequitur to sceptical arguments about free will we should not necessarily expect promoting either a deterministic world-view or a sceptical attitude to have disastrous psychological and social consequences.

The second gap in the literature is brought about by the fact that, despite their best efforts, researchers have not been able to reliably establish the nature of the relationship between belief in free will and belief in determinism. According to the traditional way the free will debate has often been framed, especially by incompatibilists, the more determined we are, the less free we are (and vice versa). But given the available evidence, of lack thereof, these distinct factors are not even negatively correlated as one might minimally expect. Rather, the work on both the FAD-Plus and the FWI suggest people's intuitions about free will and determinism are driven by largely independent psychological processes whereby the strength of belief in one does not influence belief in the other. But where the most commonly used global measures of free will have failed to find a relationship between free will and determinism, Nadlehofer et al (2014) have discovered a very small inverse correlation between determinism and a single probe designed to tap lay belief in libertarian contra-causal free will. This important finding suggests that lay beliefs about free will contain an incompatibilist element, although, more empirical spadework is required before this discovery can be claimed with any

confidence. The next chapter critiques the two psychometric tools that were used in the earlier attempts to find the real-world relationship between belief in free will and belief in determinism and finds them both to be methodologically wanting.

### 3. THE FREE WILL AND DETERMINISM SCALE

The most commonly used measurement tool in psychological studies examining effects of free will belief on behaviour is the Free Will and Determinism Scale (FAD-Plus) (Paulhus & Carey, 2011). The scale aims to assess individual differences in belief in free will and determinism on four dimensions via self-report. Participants are asked to rate twenty-seven items on a five-point Likert scale anchored at “strongly agree” and “strongly disagree.” Items on the Free Will subscale measure intuitions about free will and moral responsibility (e.g., “people have complete control over the decisions they make”); Items on the Scientific Determinism subscale measure intuitions about the biological and environmental causes of human behaviour (e.g., “as with other animals, human behaviour always follows the laws of nature”); Items on the Fatalistic Determinism subscale measure intuitions about the inevitability of the future (e.g., “Whether people like it or not, mysterious forces seem to move their lives”); and items on the Unpredictability subscale measure intuitions about luck and chance events (e.g., “Luck plays a big role in people’s lives”). With over ninety citations in published articles, the FAD-Plus has informed much of the current literature in contemporary free will research in social psychology.

The FAD-Plus is an updated version of the unpublished FAD-4 (Paulhus & Margesson, 1994) that was originally developed in response to a combination of burgeoning interest in the psychology of free will belief and the inadequacy of previous scales (Stroessner & Green, 1989; Viney, Waldman & Barchilon, 1982). Factor analysis of the earlier FAD-4 showed that items measuring belief in Free Will and Scientific Determinism clustered separately on two distinct factors thereby discrediting the earlier a priori assumption that free will and determinism are opposites of a bipolar scale (Rakos et al., 2008; Viney et al., 1982). The FAD-4 served as a useful tool in a number of published studies (Vohs & Schooler, 2008; Baumeister et al., 2009; Crescioni., et al,

2016; Stillman et al., 2010) until it was later revised and updated (Paulhus & Carey, 2011).

Despite being the measurement tool of choice in studies examining the behavioural effects of free will belief the FAD-4 had never been published as a reliable instrument due to notable psychometric weaknesses. For example, subscale intercorrelations were below the acceptable range and several items displayed double factor loadings and even cross loadings. Consequently, for almost two decades uncertainty remained about the true relations between believing in free will and determinism, and other constructs on the FAD-4. The revised version of the scale, the FAD-Plus, is believed to have eliminated many of the earlier problems inherent in the FAD-4 and continues to be the primary measure of belief in free will and determinism to this day. One advantage of the FAD-Plus is that it has separate subscales for free will and determinism which allow for the two constructs to be compared with other psychological measures (e.g., belief in free will is positively correlated with religiosity, whereas belief in scientific determinism displays a negative relation). However, despite being an improvement on the FAD-4 and contributing to almost all of the psychometric research on free will belief over the past decade the FAD-Plus has noteworthy weaknesses.

One of the major difficulties of operationalising belief in free will and belief in determinism into measurable psychometric constructs is the lack of shared vocabulary among researchers across disciplines. As the result of a long and deep engagement with the problem of free will, philosophers have a highly developed conceptual and lexical framework for describing various complex terms of the debate. For example, philosophers typically have one thing in mind when they discuss determinism in the context of free will, namely, the thesis that given the actual past and the laws of nature there is only one possible future at any moment in time. Importantly, the classical problem of free will and determinism relies on this definition to expose the inherent contradiction between the ability to do otherwise, or one's being the ultimate source of action, and the nature of cause and effect. Only determinism framed in this way engenders the critical issue of whether free will is compatible with the structure of causation.

Nadelhoffer et al (2014) write, that looking back on the earlier empirical work on free will belief, what researchers find can depend a great deal on how they conceptualise and think about free will. Not only can researcher's own beliefs influence



how they phrase questions and make predictions, but personal beliefs can also determine how results are analysed and interpreted. They call this the *problem of theory contamination* (i.e., when a researcher's own theoretical commitments unduly influence or bias their findings). If researchers are not sufficiently educated about the classical problem of free will and determinism, not only will they fail to design theoretically accurate scale constructs with the ability to measure the corresponding beliefs, they may also hold incorrect biases about what it means not to have free will. This is evident in the way Vohs and Schooler, and Baumeister and colleagues default to a fatalistic interpretation of determinism rather than demonstrating a core understanding of the sceptical view. Thus, lack of familiarity with the key terms of the debate would explain why, as newly discovered constructs in social psychology, free will and determinism have often been defined in ways discordant with philosophical definitions. This is evident in the FAD-Plus. Here Paulhus and Carey (2013, p. 130) make it clear that examining lay beliefs about free will does not necessitate consideration of rigorous philosophical definitions:

Philosophical discussions of issues regarding the concepts of free will and determinism hinge on rigorous definitions. Unfortunately, the diversity and complexity of definitions are less than helpful in the context of our empirical goals. More relevant to our goal of studying lay beliefs are bottom-up definitions. Therefore, we follow the path of those investigating folk beliefs, that is, the everyday phenomenology of free will.

But despite being determined not to get caught up in the philosophical complexities of the free will debate, like many free will researchers before them Carey and Paulhus are explicit in their goal to discover whether ordinary people hold compatibilist or incompatibilist beliefs. They write:

The research in the present report exploits the FAD-Plus to tackle the compatibility issue at the individual difference level. Preliminary work suggests that beliefs in free will and determinism are relatively uncorrelated (Paulhus & Carey, 2011, Study 2). One goal of the present article is to confirm this pattern of belief compatibility in a variety of samples (2013, p. 131).

Carey and Paulhus set out to confirm that the lay-position is compatibilism. However, the FAD-Plus does not provide any obvious way to make this distinction. One

could look for individuals with higher scores for Free Will in conjunction with higher scores for Scientific Determinism as a proxy for compatibilism. However, this does not factor in the complexity of compatibilist beliefs as it mistakenly assumes that compatibilists must believe strongly in determinism, which need not be the case. Equally, compatibilists do not have to agree with statements on the Scientific Determinism subscale that claim actions are determined by genes (Nadelhoffer et al., 2014). Consequently, trying to distinguish compatibilist from incompatibilist views by comparing average subscale scores from the FAD-Plus is problematic.

There are additional problems with the Free Will subscale. Take for example the two statements below:

- Strength of mind can always overcome the body's desires.
- People can overcome any obstacle if they truly want to.

Neither statement necessarily implies free will. Both, sentences could be taken to describe behavioural traits formed from deterministic processes such that the resulting ability to overcome the body's desires or life's obstacles is completely a consequence of prior circumstances. If free will is the ability to make decisions and perform actions that are partially or completely independent of antecedent causes, then these statements do not necessarily describe any such thing.

If compatibilism was the variable of interest items should describe free will under certain causal or limiting conditions. Statements should be worded in such a way that they tap the extent to which people attribute free will to themselves and others when there is a causal explanation for behaviour. Take for example the sentence "criminals are totally responsible for the bad things they do." If a causal factor was added (e.g., even if a bad up-bringing is the cause of their behaviour) the statement would be more likely to measure compatibilist-like beliefs. However, items on the FAD-Plus Free Will subscale do not include any such reference to causal determinants of behaviour. Furthermore, people who believe in free will do not necessarily need to believe that strength of mind *always* prevails over temptation, that *any* obstacle can be overcome, or that people are *always* at fault, and so on. The misplaced modifiers only add to the extant noise.

The efficacy of a Free Will scale relies on the statements being written in such a way that they cannot reasonably be interpreted as being about anything other than beliefs about free will. Items on the FAD-Plus do not clearly describe libertarian contra-causal free will, however, nor do they clearly indicate belief in compatibilist-like agency. Therefore, they do not appear to be obviously measuring belief in free will. Worse still, what counts as determinism in these measures only describe one or the other of the ideas that people's character and actions are caused by their genes or by their upbringing:

- Your genes determine your future.
- Parents character will determine the character of their children.
- Childhood environment will determine your success as an adult.

Not only do these items fail to describe the concept of determinism, they are also false. It requires *both* genes and environment to produce phenotypic outcomes. So, to be consistent with current empirical knowledge, scientifically minded people who would normally endorse determinism or statements describing the causal nature of gene x environment interactions may give confused ratings. With five out of seven items phrased in this way the Scientific Determinism subscale fails to capture the concept of determinism that would be required for Carey and Paulhus to tackle the compatibility issue.

In sum, poor construct validity of the item statements on the FAD-Plus may be an alternative explanation for why the researchers failed to find an association between belief in free will and belief in determinism. Understandably Paulhus and Carey's final note that their work on the FAD-Plus "supports their conclusion that lay judges see free will and determinism as quite compatible" has drawn philosophical criticism (2011, p. 102). Nadelhoffer et al (2014) write that whether Paulhus and Carey are right that their work on the FAD-Plus provides conclusive evidence that compatibilism is the default view among lay people remains an open question and that they prefer to remain agnostic for present purposes.

One potential way to distinguish lay-beliefs about (in)compatibilism is to design a scale with items that tap compatibilism and incompatibilism separately. So, in their

attempt to distinguish (in)compatibilist beliefs about free will Nadelhoffer et al. (2014) developed a more nuanced measurement tool with separate measures for both.

### 3.1 THE FREE WILL INVENTORY

Over two years Nadelhoffer et al. (2014) developed and validated the Free Will Inventory (FWI). The FWI is a two-part twenty-nine item instrument set on a 7-point Likert scale and anchored at “strongly-disagree” and “strongly-agree.” Part 1 consists of three five-item subscales designed to measure *strength* of belief in free will (e.g., “people always have free will”), determinism (e.g., “Everything that has ever happened had to happen precisely as it did, given what happened before”), and dualism/anti-reductionism (e.g., “The human mind cannot simply be reduced to the brain”). Part 2 contains 14 single item probes designed to tap facets of libertarianism, compatibilism, and moral responsibility. As a relatively new measurement tool the FWI has only recently started appearing in free will literature (Black, 2016; Genschow, Rigoni, & Brass, 2017; Monroe, Brady, & Malle, 2017).

The overarching goal of Nadelhoffer et al. was to replicate and build on the previous findings of Paulhus and Carey (2011) using a psychometric instrument with greater validity and the ability to yield more fine-grain data about relationships between free will, determinism, dualism, and facets of political psychology. They reasoned that if believing in free will goes hand in hand with worrisome beliefs and attitudes associated with intolerance, victim blaming, and belief in just deserts, a link between this disturbing constellation of behaviours and believing in free will was worthy of further investigation. Not only was it important to shed more light on the implications of believing in free will, but findings may also challenge Smilansky’s illusionist doctrine that free will belief is an instrumentally beneficial illusion worth hanging onto. Initial validation studies of the FWI supported Paulhus and Carey’s FAD-Plus finding that belief in free will and determinism are independent constructs. But to their surprise they also replicated the puzzling finding that belief in free will and belief in determinism are completely unrelated.

Consequently, Nadelhoffer et al (2014) realised if they were going to be able to distinguish compatibilist beliefs from incompatibilist beliefs and examine relationships

between other variables related to free will belief they would need to do more than only measure *strength* of belief. As the subscales in Part 1 of the FWI only measure strength of belief, the researchers knew they needed to design an additional list of items aimed at probing belief about contra-causal agency and compatibilism.

The most noteworthy finding to emerge from Part 2 was that given the chance a strong majority of people are willing to endorse statements that explicitly express libertarian views. This suggests that the sceptics are right in thinking that people ordinarily think that having free will requires contra-causal agency. Nadelhoffer et al (2014) concluded that while there is a clearly compatibilist element to how people ordinarily think about free will, it's becoming increasingly clear there is also an incompatibilist element to people's free will beliefs. However, it was hard to reconcile the finding of widespread endorsement of libertarian beliefs with the fact that both the FAD-Plus and the FWI failed to find a relationship between free will and determinism.

In a follow-up study based on Paulhus and Carey's findings linking belief in free will to various moral and political worldviews Nadelhoffer and Tochetto (2013) predicted higher scores on the FWI Free Will subscale would be associated with higher scores on right wing authoritarianism and just world beliefs, as well as several other measures. And unlike their global measure of free will belief that was unrelated to determinism, they predicted libertarian beliefs studied alone would be inversely related to belief in determinism. While, their results generally supported Carey and Paulhus's worldview findings, contrary to prediction, libertarianism was *not* negatively associated with determinism which, given the stark contradiction between the two beliefs adds to the ongoing perplexity about how people reconcile these two logically inconsistent ideas.

A second source of concern in the follow-up study is the way libertarian items were analysed. Because Part 2 of the FWI is not a valid scale, there's no grounds to suppose items grouped under various sub-headings (e.g., libertarianism, compatibilism, moral responsibility) are measuring the same construct, or even related constructs. As Nadelhoffer et al. didn't report any attempt to validate part 2 there is no way of knowing whether the fourteen probes were failed items from the exploratory rounds, or if from the outset items were specifically designed to be used as single probes. Either way, it's important to note that "libertarian items" are, for all intents and purposes, structurally unrelated. This means aggregating scores and averaging out values for comparison with

other scales is methodologically unsound. Scores yielded from the items in Part 2 should be compared with scores from the different scales and each of these correlations should be treated as a discrete result. This is precisely what Nadelhoffer et al. did in their validation study. All items from Part 2 were individually compared with the three subscales in Part 1 and results were reported individually.

One of the more notable findings was that libertarian item-1 (L1) was the only item to both correlate positively with global free will belief (as one would expect) and negatively with determinism. Given previous research on both the FAD-Plus and the FWI has failed to find a relationship between belief in free will and determinism this is a striking and informative result. So, while the finding that most people agree with statements expressing contra-causal agency suggest further investigation is warranted, as it stands, a discrete correlation yielded from a single item showing a trivial effect ( $r = -.14$ ,  $p < .01$ ) is “currently the best available empirical tool we have for discriminating libertarians from compatibilists” (Nadelhoffer & Tochetto, 2013, p. 140).

Nadelhoffer et al. predicted their new-found libertarian measure would display an inverse relation with determinism and demonstrate stronger links with different moral and political measures compared to the Free Will subscale. However, after finding a significant association between L1 and determinism the authors aggregated scores with a second libertarian item that had previously displayed a non-significant relation with determinism. Why did they do this? The authors gave no explanation for doing such. Unsurprisingly, the aggregated libertarian scores were unrelated to determinism. Considering the authors were hoping to find an inverse relation it was an unusual move to combine the item scores from Part 2.

Nadelhoffer et al. also compared four items from Part 2 with the free will and determinism subscales from Part 1. As with L1, the three items measuring compatibilism were positively associated with free will, but showed stronger inverse relations with determinism. This suggests that people who believe more strongly in compatibilism believe less in determinism and vice versa. Theoretically this is a puzzling result: given that compatibilists do not take determinism to rule out free will it is strange that compatibilists would believe less in determinism than libertarians. Furthermore, the only item in Part 2 that was designed to measure agreement with a *traditional* compatibilist conception of free will was not associated with global free will. Considering the term

“free will” is written into the statement, it’s hard to explain why traditional compatibilism and free will were unrelated. As single studies do not provide conclusive evidence further investigation of the relationships between compatibilism, free will, and determinism is needed to unpack some of these odd findings.

In sum, in their efforts to distinguish compatibilist beliefs from incompatibilist beliefs using single item probes Nadelhoffer et al (2014) yielded some surprising results. Compatibilism showed a stronger inverse association with determinism than libertarianism, and the only statement describing traditional compatibilism was unrelated to a global measure of belief in free will. Although the authors didn’t offer explanations for these puzzling results, they may be due to unclearly defined terms tapping intuitions that overlap with libertarian beliefs and items of the global free will scale not accurately describing free will. To that end, with respect to distinguishing compatibilist beliefs from incompatibilist beliefs and exploring their relative associations the FWI did not live up to expectations.

### 3.2 SCALE DEVELOPMENT

So, while compatibilism and incompatibilism are *logically* incompatible, uncertainty remains about whether and to what extent they are *psychologically* incompatible. Their psychological (in)compatibility suggests that working from the assumption of a unified model of free will is perhaps not going to yield the most accurate and therefore meaningful results. As such, to have a better chance of finding the elusive relationship between belief in determinism and belief in free will, the first task of the current study was to develop a validate a scale that described these distinct facets of free will belief independently.

Correlational studies that examine links between belief in free will and various attitudes, behaviours and beliefs currently employ global measures (FAD-Plus, FWI) that assume a unified model of free will (Baumiester & Brewer, 2012; Carey & Paulhus, 2013; Genschow et al., 2017; Nadelhoffer & Tocchetto, 2013). However, findings from psychometric and other survey literature show lay theories tend to be a mix of both compatibilism and elements of incompatibilism. For example, Monroe and Malle (2014) found that when asked directly, most people’s core concept of free will entails a choice

that fulfils one's desires and is free from internal or external constraints. No evidence was found for metaphysical assumptions about dualism or indeterminism. Conversely, Nadelhoffer et al. (2014) discovered that, when given the opportunity most people are willing to endorse statements that explicitly express libertarian contra-causal views. This latter finding is tremendously important because it shows that belief in contra-causal agency can be examined on its own. The current study extends the findings of Nadelhoffer et al. (2014) by designing a measurement tool with the ability to measure belief in contra-causal free will and belief in compatibilist-like agency as two distinct facets of belief in free will. I will refer to it as the *Free Will Scale* (FWS).

Items on the FWS are generally modified statements from previous scales. For example, because the critical L1 probe in Nadelhoffer et al. (2014) was a description of free will that in theory both libertarians *and* free will sceptics could agree to (i.e., "Free will is the ability to make different choices even if everything leading up to one's choice [e.g., the past, the situation, and their desires and mental state etc.] were exactly the same") it was changed to, "People have the ability to make different choices even if everything leading up to their choice, including their mental state, was exactly the same." Furthermore, several of the confusing modifiers were removed, for example: "People *always* have the ability to do otherwise" (Nadelhoffer et al., 2014, p. 34) was changed to "A person could have, of their own accord, done otherwise than they did." Importantly, items on the contra-causal agency subscale and items on the compatibilist-like subscale were worded in such a way that a person holding one view could disagree with all the items describing the other view and vice versa.

### 3.3 PREDICTIONS

Based on the research mentioned above, I made the following predictions:

- Scores on the libertarian contra-causal subscale will be inversely related to scores on the determinism subscale. Whereas, scores on the compatibilist-like subscale will be unrelated to determinism.
- Right-wing authoritarianism and belief in a just world will be positively related to belief in free will on both dimensions.



- Compassion, gratitude, and satisfaction with life will not be related to belief in free will and determinism.
- Sceptics will score lower than non-sceptics on the right-wing authoritarianism, belief in a just world, and retribution.
- Non-sceptics will score higher on internal locus of control, and sceptics will score higher on external locus of control
- There will be no differences between sceptics and non-sceptics for gratitude, and satisfaction with life.
- Sceptics will score higher on compassion.

## 4. METHOD

### *4.1 Participants*

A cluster sample of lay-people who hold positions in the free will debate ( $n = 92$ ) were recruited via three public philosophy discussion groups, two on Facebook and one on Reddit. Half the sample was American and half were a mixture of Europeans, Canadians, Australians and New Zealanders. A sample of Americans ( $n = 191$ ) were recruited via Amazons Mechanical Turk, and a European/Australasian sample ( $n = 113$ ) self-selected from various general Reddit pages by country. The total sample ( $n = 396$ ) were 70% male, ages ranged from 18-55+. The Mturk group was significantly more religious and less liberal than the Facebook and the Reddit groups. The total sample yielded ( $n = 77$ ) sceptics, ( $n = 37$ ) Compatibilists, ( $n = 39$ ) Libertarians, and ( $n = 242$ ) people with no view in the debate.

### *4.2 Materials and procedure*

Participants completed an online survey powered by Qualtrics. They were informed that the survey was anonymous, that ethics approval from Victoria University Human Ethics Committee had been granted, and that completion of the survey served as informed consent for use of their responses in this research. The American MTurk participants were each paid a small fee (USD \$1.50) for their participation in the study.

All psychometric questionnaires used in this study were set on a 7-point Likert scale anchored at 1-Strongly disagree and 7-Strongly agree.

*The Free Will Scale.* Items on the free will scales have been modified from previous scales (Paulhus & Carey, 2011., Nadelhoffer et al, 2014., Derry, Davis & Carey, 2014) and developed to measure belief in free will on two dimensions – belief in *contra-causal agency* and *agreement with compatibilist-like agency*. Both subscales displayed good alphas ( $\alpha = .90$ ,  $\alpha = .88$ ) respectively.<sup>15</sup>

*Items on the Free Will Scale.*

|   | <i>M</i> | <i>SD</i> |
|---|----------|-----------|
| <b><i>Belief in contra-causal agency</i></b>  |          |           |
| People have the ability to make different choices even if everything leading up to their choice, including their mental state, was exactly the same.        | 4.31     | 2.0       |
| A person could have, of their own accord, done otherwise than they did.   | 4.47     | 1.97      |
| People can override their genetic inheritance and life experience when making important decisions.  | 4.30     | 1.88      |
| Our past doesn't always affect our future because we have ultimate control over our actions.  | 4.05     | 2.06      |
| My genetic make-up and past experiences do not always influence my decisions and actions.   | 3.99     | 1.97      |
| <b><i>Agreement with Compatibilist-like agency</i></b>  |          |           |
| As long as your decisions are based on your own personal beliefs and values, that's enough by itself to be blameworthy or praiseworthy for those decisions. | 4.74     | 1.71      |
| Even though I'm not responsible for all the influences that shaped me, as long as I act on my personal convictions I'm exercising my free will.             | 4.72     | 1.90      |
| As long as a person is of sound mind and not being coerced, that's enough to say they are exercising their free will.                                       | 4.89     | 1.84      |
| Criminals should be blamed for the bad things they do even if bad up-bringing is the cause of their behaviour.  | 4.56     | 1.99      |
| Although there are always situational and personal reasons someone acts the way they do, they are ultimately responsible for how they behave.               | 4.44     | 1.91      |

*Determinism Subscale* (FWI: Nadelhoffer et al, 2014). The determinism subscale is a 5-item scale developed to measure belief in philosophical determinism. Items include “Everything thing that has ever happened had to happen precisely as it did, given what happened before” and “given the way things were at the big bang, there is only one way for everything in the universe to happen after that” ( $\alpha = .90$ ).

<sup>15</sup> Cronbachs alpha is a measure of internal consistency. That is, how closely related a set of items are as a group. It assumes that the average correlation of a set of items is an accurate estimate of the reliability of a psychometric scale. Alphas above .70 are generally considered acceptable (Tavakol & Dennick, 2011).

*Right-wing Authoritarian scale* (RWA; Zarkrisson, 2005). The RWA Scale was developed to measure the degree of obedience to authority and adherence to existing social conventions. Scores on RWA have been found to reliably correlate with attitudes of punitiveness, racial prejudice, homophobia, religious orthodoxy, and victim blaming. The shortened version of the RWA consists of five items that include “Our society would be better off if we showed tolerance and understanding for untraditional values and opinions” and “People ought to put less attention on the Bible and religion, instead they ought to develop their own moral standards” (items are reverse coded), ( $\alpha = .85$ ).

*Belief in a Just World Scale* (JWB; Lambert, Burroughs & Nguyen, 1999). The JWB Scale measures the degree to which individuals believe they live in a world where people generally get what they deserve. People who score highly on JWB are more likely to trust current institutions and authorities and to blame the poor and praise the rich for their respective fates (Lambert, et al, 1999). The shortened version contains six items reflecting just world attitudes towards *others* (e.g., “I believe that, by and large, people get what they deserve”) and six items reflecting just world attitudes about the *self* (e.g., “I believe that I usually get what I deserve.”), ( $\alpha = .84$ ,  $\alpha = .83$ ) respectively.

*Retribution Scale* (Gerber & Jackson, 2013). Retribution is a six-item questionnaire (the Retribution as Revenge Subscale) that measures attitudes about punishment in the context of criminal justice. Gerber and Jackson (2013) have shown that attitudes towards the use of harsh punishments to get even with the offender go hand in hand with the idea that punishment should make the offender suffer. Participants rate their strength of agreement with retributive punishment goals on two dimensions – *make suffer* (e.g., “Punishment is deserved suffering”) and *get even* (e.g., “We should punish to get even with the offender”), ( $\alpha = .91$ ).

*Treatment of Criminals Scale* (Gerber & Jackson, 2013). The scale measures people’s beliefs about how criminal offenders should be treated on two dimensions – harsh punishment and procedural fairness. Four items (two reverse coded) measure support for punitive measures (e.g., “People who break the law should be given harsher sentences”)

and three items (two reversed) measure attitudes about the offenders right to dignity and respect (e.g., “After committing an offence, criminal offenders lose the right to be treated with respect”), ( $\alpha = .87$ ).

*Compassion Scale* (Pommier & Neff, 2010). The *Common humanity* subscale is an 8-item subscale from the Compassion Scale that measures dispositional compassion towards suffering. Sample items include “Suffering is just a part of the common human experience” and “I can’t really connect with other people when they’re suffering” (reverse coded), ( $\alpha = .78$ ).

*The Gratitude Questionnaire-6* (GQ6; McCullough, Emmons & Tsang, 2002). This is a six-item questionnaire assessing gratitude. Higher scores on the GQ6 are positively associated with positive affect and well-being, prosocial behaviours, and religiosity. Items include “I have so much in life to be thankful for,” and two reversed items (e.g., “When I look at the world I don’t see much to be grateful for”), ( $\alpha = .84$ ).

*Satisfaction with Life Scale* (SWLS; Diener, Emmons, Larson & Griffin, 1985). The SWLS is a five-item scale assessing general subjective life satisfaction. Scale items include “In most ways, my life is close to ideal” and “If I could live my life over, I would change almost nothing.” ( $\alpha = .81$ ).

*Brief Locus of Control Scale* (Sapp & Harrod, 1993). The scale measures perceptions of internal versus external control on a three item measure of internal control (e.g., “My life is determined by my own actions”) and two three items measures of external control *chance* (e.g., “When I get what I want it’s because I’m lucky”), and *powerful others* (e.g., “My life is chiefly controlled by powerful others”). Alphas were acceptable at ( $\alpha = .70$ ,  $\alpha = .75$ ,  $\alpha = .76$ ) respectively.

## 5. RESULTS

The first aim of this study was to establish whether the contra-casual agency subscale and compatibilist-like subscales in fact measure two different concepts.<sup>16</sup> To estimate parameters, and to confirm factor loadings and model fit a confirmatory factor analysis (CFA) with belief in contra-causal agency, agreement with compatibilist-like agency, and agreement with determinism<sup>17</sup> as latent factors was carried out in AMOS. Little's MCAR test showed missing data was Missing Completely at Random and was imputed for each of the 16 scales using Expected Maximisation in SPSS. Regression weights were estimated using the maximum likelihood method and bootstrapped at 2000 iterations. Chi-square, Comparative Fit Index (CFI), the Root Mean Square Error of Approximation (RMSEA), Standardised Root Mean Square Residual (sRMR), and HOELTER's Critical N were used to test model fit. Chi-square and the Comparative Fit Index (CFI) compares the similarity of the predicted model covariance matrix and the observed covariance matrix; non-significant Chi-square values below 5.0 and CFI's above .95 are desirable. RMSEA indicates the likelihood of the model being a close fit to the population, and values of .05 are optimal. The sRMR is the average residual correlation left over after fitting the model to the data and should be under .05. Critical N estimates the sample size adequate for obtaining a good model fit for the chi-square test and these values should be greater than 200.

All model fit indices were within acceptable ranges however covariance between the two latent structures was unacceptably strong indicating multicollinearity ( $r = .91$ ,  $p$

---

<sup>16</sup> A pilot confirmatory factor analysis ( $n = 97$ ) using AMOS showed good model fit across all fit indices. However, there was an unacceptably high correlation between belief in contra-causal agency and agreement with compatibilist-like agency  $r(97) = .76$ ,  $p < .001$  indicating the subscales were not tapping differential responses between the two conceptions of free will. The compatibilist-like subscales were revised and the language was strengthened via the inclusion of the "free will" term and elements suggesting desert based moral responsibility.

<sup>17</sup> I elected to include the Determinism subscale from the FWI for three reasons: first, the initial CFA displayed a low CFI (.93) (Nadelhoffer et al, 2014). Second, as a relatively new measure the determinism subscale did not have a consistent track record of construct validity. Third, it had not shown an inverse association with a valid multiple-item measure of free will.

= .001). This meant participants were responding to the two conceptions of free will too similarly to separate them for a structural equation analysis. However, examination of the covariance within groups revealed a more complex story. While the Mturk group displayed a near perfect score ( $r = .99, p = .001$ ), the Reddit group was well within the acceptable range for justification of a two-factor model of free will ( $r = .55, p = .001$ ). Moreover, the model fit indices of a CFA conducted only in the Reddit group indicated a superior model fit than that of the whole sample. Nevertheless, for present purposes the model was modified into a two factor structure with global free will and determinism as the latent variables.

The optimal number of observed variables to model onto a latent structure is between five and two with three being optimal. The Free Will alpha indicated excellent internal consistency, thus providing justification to parcel items to optimise the number of observed variables. Parcelling is the systematic process of combining items into pairs and triplets then averaging scores for the purpose of decreasing numbers of observed variables. In the current study, the five items on each subscale were parcelled to produce two new items per subscale that could then be tested against a theoretical model with belief in global free will and belief in determinism as the latent factors.

All items loaded highly onto the two factors (See fig 1. for standardised coefficients) and all were significant. The model fit indices showed that the CFI (.98) was good, RMSEA was reasonable .07 within 95% CI's (.058, .094),  $p = .009$ , sRMR (.038) was below .05, and Hoelters Critical N values was above 200 (212). Chi-square was significant  $\chi^2(25) = 2.127, p = .001$ . In general, larger samples will mostly cause the Chi-square statistic to display a significant difference between the model estimations and observed data suggesting misfit irrespective of the true situation (Byrne, 2010). To that end, the significant chi-square in this case is most likely a result of a large sample size ( $n = 396$ ). Taken together, although RMSEA was on the high side model fit indices indicated a fair model fit in this sample.

Figure 1

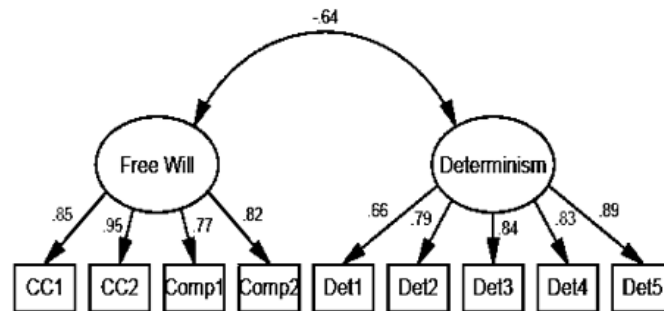


Fig 1. Model displays standardised coefficients

### 5.1 Discussion

The most important finding to emerge from the CFA is a strong inverse association between free will and determinism (fig 1.). This result provides preliminary support for the stated prediction, namely that relatively greater belief in determinism is associated with decreased belief in free will and vice versa. While this result confirms earlier findings that show free will and determinism items reliably load onto independent factors (Paulhus & Carey, 2011; Nadelhoffer et al, 2014), the discovery of a strong inverse relation between free will and determinism contradicts empirical findings to date that have shown the two constructs to be unrelated. Thus, the current finding breaks new ground within the free will literature.

With respect to establishing the two conceptions of free will as separable dimensions, although it's disappointing to find that belief in contra-causal agency and agreement with compatibilist-like agency is subject multicollinearity in this sample, when analysed separately, the largely European Reddit group displayed a factor coefficient that was within the acceptable range to justify independent analysis. This meant the Reddit subjects differentially responded to statements describing contra-causal agency and compatibilist-like agency, whereas the MTurk group did not. This suggests that Mturk workers either do not make a theoretical distinction between contra-causal agency and compatibilist-like agency, or perhaps the two dimensions of free will belief do not generally operate as independent cognitive processes in this population.

## 5.2 Correlations and construct validity of a two dimension model of free will

One of the main aims of the current study is to test the links between belief in free will and determinism, and moral beliefs and attitudes (Paulhus & Carey, 2011; Nadelhoffer et al, 2014). Although significant, many of these associations in the previous studies were relatively weak and my hope was that a scale designed to more closely resemble key philosophical definitions would produce more fine-grained data with stronger effects. Partial correlations controlling for religiosity and political orientation were performed between belief in free will, belief in Determinism and right-wing authoritarianism, belief in a just world, retribution, compassion, gratitude, and satisfaction with life.<sup>18</sup> Given the high covariance between belief in contra-causal agency and belief in compatibilist-like agency in this sample, predictably the two dimensions of free will displayed similar correlation values across all twelve variables. With the exception of determinism, I elected to combine the subscale scores and report them as a global free will.

**Table 1..** Correlations between believing in free will and determinism, and outcome variables (*Df* = 391)

| Variable                           | Free will | Determinism | <i>M</i> | <i>SD</i> |
|------------------------------------|-----------|-------------|----------|-----------|
|                                    | <i>r</i>  | <i>r</i>    |          |           |
| <i>Free will</i>                   | —         | —           | 4.56     | 1.44      |
| <i>Determinism</i>                 | -.64**    | —           | 3.85     | 1.65      |
| <i>Contra-causal agency</i>        | .95       | -.65**      | 4.34     | 1.59      |
| <i>Compatibilist-like agency</i>   | .93       | -.54**      | 4.78     | 1.45      |
| <i>Right wing authoritarianism</i> | -.08      | .08         | 2.57     | 1.08      |
| <i>Others JWB</i>                  | .42**     | -.01        | 3.76     | 1.24      |
| <i>Self JWB</i>                    | .45**     | -.16        | 4.59     | 1.06      |
| <i>Retribution</i>                 | .45**     | -.03        | 3.39     | 1.51      |
| <i>Treat of criminals</i>          | .22**     | -.01        | 3.07     | 1.34      |
| <i>Compassion</i>                  | .01       | -.13        | 5.38     | .89       |
| <i>Gratitude</i>                   | .16       | -.13        | 5.18     | 1.16      |
| <i>Sat with life</i>               | .15       | -.02        | 4.16     | 1.35      |

Note, *N* = 395. The table displays Pearson correlation coefficients, means, and standard deviations. Scales were anchored at 1-Strongly disagree to 7-Strongly agree. \*\* = *p* < .0019.

<sup>18</sup> The family-wise error rate was calculated for 27 correlations and the alpha level was adjusted to *p* < .0019. Correlation values of < .20 account for less than 5% of the shared variance between variables and are therefore of no practical importance, .20 – .30 indicate weak associations, .30 – .50 moderate associations, .50 – .70 strong associations, and .70 + very strong associations (McGrath & Meyer, 2006).



As predicted belief in *contra-causal agency* shows a strong inverse association with determinism indicating that where the contra-causal assumption is strongest belief in determinism is weakest and vice-versa. To rule out the possibility that the effect reached significance as a result of including free will sceptics in the sample – sceptics mostly believe in determinism and disbelieve in contra-causal agency – analysis of the correlational patterns within groups by source was carried out. Results show the effect is significant across groups, Facebook ( $r = -.78, p = .001$ ), Reddit ( $r = -.63, p = .001$ ), Mturk ( $r = -.30, p = .001$ ).<sup>19</sup>

Results were more complicated for scores on *compatibilist-like agency*. While there was a strong inverse association between compatibilist-like statements and determinism across the whole sample, examination of the correlations within groups revealed significant associations for the Facebook group ( $r = -.59, p = .001$ ) and Mturk group ( $r = -.35, p < .001$ ). However, the Reddit group failed to show the effect ( $r = -.09, p = .749$ ).

There was no association between right-wing-authoritarianism and either free will or determinism. On the other hand, believing in free will was moderately positively correlated with *belief in a just world* for both *self* and *others*, and *retribution*, and weakly related to the *treatment of criminals*. Those with stronger belief in free will are more likely to believe that people get their just deserts in life, that society has the right to get even with criminals and that suffering is an implicit part of the punishment. Finally, as predicted, *compassion*, *gratitude*, and *satisfaction with life* were unrelated to belief in free will and belief in determinism.

### 5.3 Discussion

This correlational study replicated several previous findings and provided construct validity for a two-factor model of free will beliefs. While the inverse relationship between belief in contra-causal agency and belief in determinism was consistent with theoretical expectations and robust across groups, my prediction that

---

<sup>19</sup> Correlations between determinism and the two dimensions of free will are taken from the structural equation model. All other values are calculated from the means.

belief in compatibilist-like agency would be unrelated to determinism was only shown in the Reddit group. Not only does this result suggest an interesting cross-cultural difference between Americans and people from other Western countries, but it is of tremendous theoretical interest that these two dimensions of free will show differential interactions with belief in determinism. Critically, it provides tentative empirical support for the idea that the dimensions are separable psychological mechanisms that potentially serve independent functions.

After controlling for religion and political orientation I found no relationship between belief in free will or determinism and right-wing authoritarianism. However, while this relationship was significant in previous studies the effect was generally weak. I used a shortened reverse coded 5-item version of the scale (Zarkrisson, 2005) rather than the 15 item version used in Carey and Paulhus (2013). It is possible that this scale was not sensitive enough to replicate previous findings. However, it was a different story for belief in a just world and measures of retribution and the treatment of criminals. The FWS replicated previous results and explained 20% of the shared variance between belief in free will and belief in a just world rather than 5% in Naddlhofer and Tochetto (2013). This lends strong support to work on both the FAD-Plus and the FWI that found that where belief in free will is strongest, so too is the belief people get their just deserts in life. Finally, results confirmed my prediction that believing in free will is unrelated to compassion, gratitude and satisfaction with life. This suggests that rather than tapping the contra casual assumption the free will subscale on the FAD-Plus may simply be tapping other positive traits that correlate with psychologically positive variables.

#### *5.4 The effects of real-world scepticism*

Another aim of the current study was to examine the beliefs and attitudes of real world free will sceptics. The priming techniques used to weaken belief in free will in previous studies have been subject to powerful criticisms; namely, that rather than priming disbelief in free will they are actually inducing a temporary feeling of fatalism that is not generalisable to people outside of these particular experimental conditions. I wanted to examine the long term psychological effects of sustained disbelief in free will so I have compared the beliefs and attitudes of people naive to the debate with people that

have taken the time learn a little bit about the topic, who do not conflate determinism with fatalism, and who are actual free will sceptics.

Figure 2

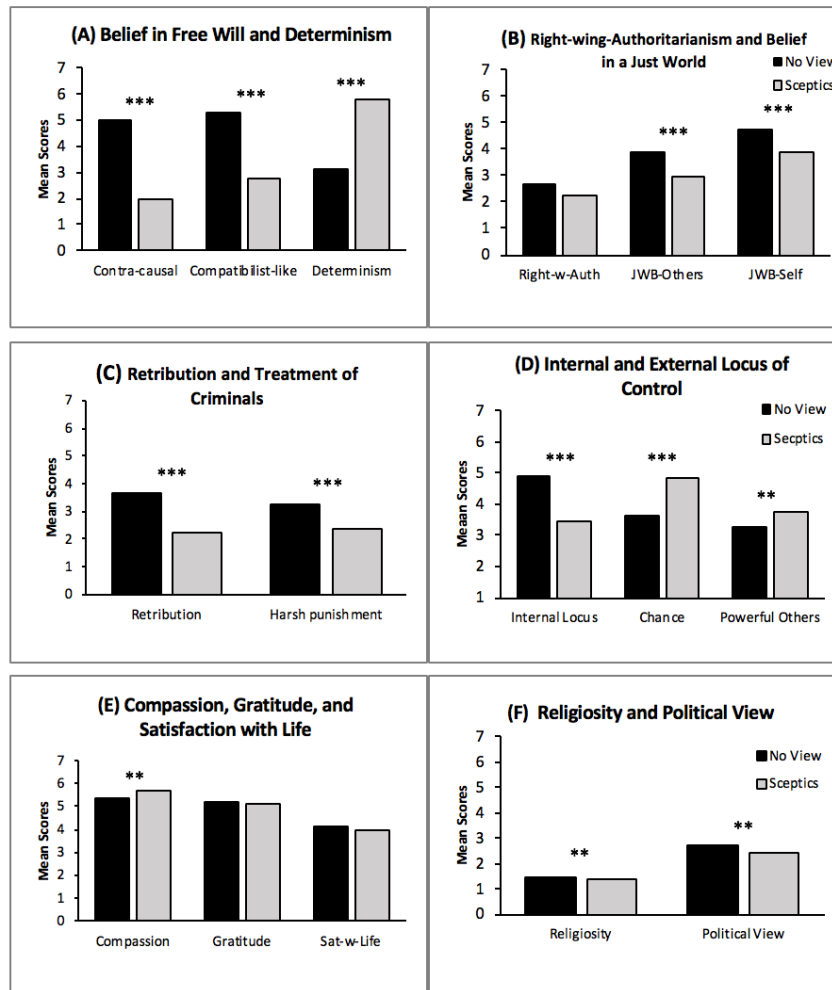


Fig 2. Bar graphs show mean scores for sceptics and participants with no view for (A) belief in contra-causal, agreement with compatibilist-like Agency, and belief in determinism. (B) Right-wing-authoritarianism, and belief in a just world for others and self. (C) agreement with retributive punishment goals and harsh treatment for criminals. (D) Internal locus of control, external locus of control chance and powerful others. (F) Religiosity an political orientation. Scales are anchored at Strongly-1 Agree to Strongly Disagree-7, religiosity was anchored at None at all-1 to A great deal-5. \*\* =  $p < .05$ , \*\*\* =  $p < .001$ .

Multivariate analysis of variance (MANOVA) was performed between groups across 16 variables. The assumption of heterogeneity was met for all but compatibilist-like agency. As predicted there was a strong effect of group on belief in contra-causal agency ( $F(1, 318) = 496.88, p < .001, \eta^2 = .61$ ), compatibilist-like agency ( $F(1, 302) = .81, p < .001, \eta^2 = .49$ ), and determinism ( $F(1, 318) = 252.86, p < .001, \eta^2 = .44$ ). Non-

sceptics believed in both conceptions of free will significantly more than sceptics, whilst sceptics believed significantly more in determinism than non-sceptics. (See table 2 for Means and SD's.) There was a strong effect of group on belief in a just world for others ( $F(1, 318) = 46.61, p < .001, \eta^2 = .13$ ), and a medium effect on belief in a just world for self ( $F(1, 318) = 32.80, p < .001, \eta^2 = .09$ ). Non-sceptics believed more strongly than sceptics that they and others get their just-deserts in life. There was no group difference for right-wing-authoritarianism. There was a strong effect of group on retribution ( $F(1, 318) = 63.64, p < .001, \eta^2 = .17$ ), and a medium effect on the treatment of criminals. Participants with no view agreed more strongly than sceptics with retributive punishment goals and harsher punishments for criminal offenders.

There were strong main effects of group on *internal locus of control* ( $F(1, 318) = 120.21, p < .001, \eta^2 = .28$ ), and *external locus of control (luck)* ( $F(1, 318) = 51.52, p < .001, \eta^2 = .14$ ), and a very small effect on *external locus of control (powerful others)* ( $F(1, 318) = 7.45, p = .007, \eta^2 = .02$ ). Non-sceptics scored significantly higher for internal locus of control whereas sceptics believed significantly more in the role of luck and the influence of powerful people in determining what happens in life.

There was a small main effect of group on compassion ( $F(1, 318) = 6.67, p = .01, \eta^2 = .02$ ). Sceptics were slightly more compassionate than non-sceptics. There were no group differences on measures of gratitude ( $p = .48$ ), and satisfaction with life ( $p = .36$ ). There was very small main effect of group on religiosity ( $F(1, 318) = 5.15, p = .024, \eta^2 = .016$ ), and political view ( $F(1, 318) = 12.27, p = .001, \eta^2 = .04$ ). Non-sceptics were slightly more religious and slightly more conservative than sceptics. However, the between group differences are very small.

□ **Table 2.** Mean scores and standard deviations for subjects with no view and free will sceptics.

| <i>Variable</i>                         | <i>No View</i> |           | <i>Sceptics</i> |           |
|---|----------------|-----------|-----------------|-----------|
|   | <i>M</i>       | <i>SD</i> | <i>M</i>        | <i>SD</i> |
| <i>Contra-causal agency</i>             | 5.0            | 1.02      | 2.0**           | 1.06      |
| <i>Compatibilist-like agency</i>        | 5.28           | .93       | 2.79**          | 1.50      |
| <i>Determinism</i>                      | 3.18           | 1.25      | 5.76**          | 1.21      |
| <i>Right wing authoritarianism</i>      | 2.66           | 1.13      | 2.23            | 1.13      |
| <i>Others JWB</i>                       | 3.90           | 1.14      | 2.93**          | .91       |
| <i>Self JWB</i>                         | 4.72           | 1.07      | 3.95**          | 1.07      |
| <i>Retribution</i>                      | 3.65           | 1.34      | 2.26**          | 1.30      |
| <i>Treat of criminals</i>               | 3.25           | 1.35      | 2.34**          | 1.18      |
| <i>Internal locus</i>                   | 4.89           | .98       | 3.45**          | 1.07      |
| <i>External locus (chance)</i>          | 3.63           | 1.26      | 4.82**          | 1.30      |
| <i>External locus (powerful others)</i> | 3.29           | 1.27      | 3.74*           | 1.32      |
| <i>Compassion</i>                       | 5.41           | .793      | 5.68*           | .893      |
| <i>Gratitude</i>                        | 5.24           | 1.16      | 5.36            | 1.37      |
| <i>Sat with life</i>                    | 4.15           | 1.35      | 3.98            | 1.36      |
| <i>Religiosity</i>                      | 1.79           | 1.19      | 1.45*           | .95       |
| <i>Political orientation</i>            | 3.10           | 1.59      | 2.39*           | 1.46      |

Note, scales are anchored at Strongly Agree to Strongly Disagree. Religiosity was set on a 5-point Likert scale anchored at None at all to A great deal. \* =  $p < .05$ , \*\* $p < .001$ . □

## 5.5 Discussion

This was the first study of its kind in free will literature to examine and compare the beliefs and attitudes of free will sceptics with people naive to the debate. The results showed strong differences between sceptics and non-sceptics on strength of belief in both conceptions of free will and strength of belief in determinism. These results give reasonable grounds for confidence that the FWS together with the FWI Determinism subscale are doing the work they are intended to do.

While there was no group difference for RWA this may be attributable to both groups being politically left leaning and the 5-item scale simply not being sufficient to capture the authoritarian trait. On the other hand, consistent with previous findings, and in accordance with the free will scepticisms of Pereboom, Waller, and Caruso, sceptics believed significantly less than non-sceptics in a just world, retribution and the harsh treatment of criminals. They were also more compassionate and emphasised the role of luck and powerful others over personal control and choice in determining life outcomes. But as with the above correlational study, unlike Crescioni et al (2016) who found those

with stronger belief in free will tend to be more grateful and have more satisfaction with life using the FAD-Plus I found no such differences.

## 5.6 GENERAL DISCUSSION

The current study was conducted with several goals in mind. First, I wanted to develop and validate a measurement tool that would more closely resemble the philosophical definitions of the key components of the debate and that would also align with the expectations of social psychologists, that what they are in fact measuring is belief in libertarian contra-causal free will (Vohs & Schooler, 2008; Baumeister, 2009; Clark et al., 2017). Second, I wanted the new measurement tool to have the ability to tap incompatibilist and compatibilist-like beliefs independently, and in doing so to hopefully find the mysteriously absent inverse relationship between the contra-causal assumption and determinism that would reveal lay incompatibilism. Third, for the first time in free will research I wanted to compare the beliefs and attitudes of real-world sceptics with people naive to the debate to show, compared to the flawed priming experiments, the sort of attitudes and behaviour we could expect from people if free will scepticism was widely accepted.

Previous correlational studies that have examined links between belief in free will, belief in determinism, and various moral and political beliefs employed measurement tools that assumed a unified model of free will that was not composed of theoretically distinct constructs for the contra-causal assumption and compatibilist-like beliefs. Considering these two beliefs are logically inconsistent it is rather surprising that a scale measuring belief in global free will would not have subscale items designed specifically to tap each of these incompatible beliefs independently. Results from the new measurement tool on two dimensions' support and extend the recent findings of Nadelhoffer et al (2014) who, using a single item probe, found preliminary evidence showing that when given the opportunity most people will endorse contra-causal libertarian views, and that this view is inversely related to determinism. People believe the more determined we are, the less free we are, and vice versa. The current study lends

strong evidential support to the Nadelhoffer et al (2014) finding that there is an incompatibilist element to people's beliefs about free will.

But as is often the case, imposing a theoretical model onto complex human beliefs out in the real world does not generally yield straight forward results and often leads to more questions than it answers. A confirmatory factor analysis with the two conceptions of free will as latent variables indicated justification for their independent analysis in a European/Australasian population but not in an American one. This unexpected cross-cultural effect suggests different groups interpret and understand statements describing the two conceptions of free will in different ways. This in turn has produced an unusual situation whereby a two-factor structure for belief in free will is valid in one population but not in another. While this presents new challenges in terms being able to distinguish incompatibilist from compatibilist beliefs using psychometric data, it does however highlight how sensitive external reliability is to the idiosyncrasies of a given sample.

I think it is important to note that there were some fundamental differences between these two samples that may explain this asymmetry of results. The MTurk workers were a paid sample who spent an average of five minutes less time on the questionnaire than their self-selected Reddit and Facebook counterparts. With the exploding use of MTurk for internet surveys, research efforts and data quality are increasingly subject to the effects of "professional survey takers" who do not give the required attention to survey questions and who speed through the survey, or who intentionally cheat with their answers (Smith et al, 2016). This was evident in the MTurk sample. The much shorter time MTurk workers took to complete the questionnaire as well as being financially motivated to participate rather than choosing to take part out of interest or to contribute to psychological research give reason to doubt the quality of the MTurk data. As such I would be reluctant to generalise the MTurk results to other settings.

On the other hand, the self-selecting Reddit sample showed that belief in libertarian contra-causal agency is inversely related to belief in determinism indicating lay-incompatibilism. In contrast compatibilist-like beliefs were unrelated to determinism indicating that the extent to which people agree with compatibilist type ideas about free will and moral responsibility does not either logically or psychologically rule out their believing in determinism. People can believe both views, neither, or one and not the other. This finding supports Nadelhoffer and Tochetto's (2013) claim that the debate about lay

incompatibilism is based on a false dichotomy because people appear to hold both compatibilist and incompatibilist views. I believe this is the first study to show a theoretically predictable three-way relationship between the belief in the contra-causal assumption, agreement with compatibilist-like agency, and belief in causal determinism. The finding that these two views about free will differentially interact with determinism suggests that they may even serve different cognitive functions that could potentially be studied independently. Future work should attempt to explore how they interact with a range of traits and attitudes in diverse populations so that a more nuanced picture can emerge.

Most of the correlational predictions were borne out by the data and support the findings of Carey and Paulhus (2013) and Nadelhoffer and Tochetto (2013) who found that where free will is strongest so is belief in just deserts and retributive punishments. Although it's difficult to establish causal direction from concurrent data, we may hypothesise from the present findings that increasing people's awareness of the effects of determinism on human agency diminishes belief in free will which in turn may reduce belief in a just world and retributive behaviour. This supports the work of Shariff et al. (2014) who showed that exposure to mechanistic explanations of human behaviour via a neuroscience class reduced blame and punitiveness towards criminals. I think it's important to mention here, that while I found the consistency of findings across the previous studies compelling, neither of the scales these researchers used to measure free will was inversely related to determinism. In light of both the present results and the logical relationship of the ideas, this suggests that that they were *not* measuring the contra-causal assumption. Furthermore, as neither scale contained specific compatibilist items, it remains unclear what exactly their scales were measuring. The current study used a scale that was reliably incompatibilist and replicated the previous findings with much stronger effects. This is the strongest evidence yet that believing in free will goes hand in hand with believing people justly deserve what they get in life and retributive attitudes.

The new scale allowed me to reject the idea that sceptics are less grateful and experience less life satisfaction than their non-sceptic counterparts (Crescioni, et al., 2016). Free will scepticism does not necessitate a lack of gratitude, and here may be why. Whilst gratitude involves being thankful to someone who has acted beneficially and being thankful may invoke a sense that the person acting beneficially is praiseworthy in the



desert-based sense, it is entirely possible to retain the feeling of being thankful without believing they are morally responsible in this way. Gratitude typically involves appreciation as a response to a kind act someone has performed towards us. My finding that there was no difference between sceptics and non-sceptics for gratitude supports the idea that the sceptical rejection of desert-based moral responsibility does not pose a threat to the legitimacy of being thankful and expressing appreciation when others are considerate. The same could be said for being satisfied with one's life. The fact that there was no group difference in SWLS shows that giving up on the idea of free will and desert-based judgements does in no way detract from the good life. The Baumeister et al (2009) finding that stronger belief in free will was related to the gamut of well-being traits is most likely explained by the fact that the FAD-Plus subscale was not measuring free will, but rather some other facet of agency related attitudes.

In the experimental studies, Vohs and her colleagues have suggested and that people with weaker belief in free will "seem to, at least temporarily, abandon their moral code" (Shariff et al., 2008, p. 198), and Baumeister and colleagues has claimed that "belief in free will may be crucial for motivating people to control their automatic impulses in favour of more prosocial forms of behaviour" (2009, p. 261). But as we have seen, it is doubtful that Vohs and Schooler (2008), and Baumeister and colleagues (2009) have shown any such thing, because what they have been studying appears to be the temporarily induced effects of a sense of fatalism. So contrary to many of these sorts of claims in the psychology literature, we appear to have seen no evidence to date that a deterministic world view or free will scepticism view encourages anti-social behaviour.

On the other hand, by studying the people who have taken the time learn a bit about the topic, who do not conflate determinism with fatalism, and who are actual real world free will sceptics I have presented evidence of the social goods we could come to expect if free will and moral desert was widely rejected. More compassion, less belief in the pernicious "just deserts," and a rejection of retributive justice in all its punitive forms. Furthermore, sceptics, compared to non-sceptics emphasised the causal role of external influences such as luck and the actions of others in shaping individual lives, over the radical self-making world-making beliefs of people with a stronger internal locus of control. The tendency of sceptics to view factors external to the self as having a greater influence than personal choice and control in determining life outcomes, supports Wallers

claim, that scepticism about free will and moral desert would enhance our ability to look more clearly at the causes and deeply at the into the systems that shape individuals and their behaviour.

## 6. CONCLUSION

Defenders of free will illusionism claim that because people's illusory beliefs about libertarian free will underpin their ordinary notions of desert-based moral responsibility, if we are to prevent society from descending into moral nihilism then these beliefs should be preserved as much as possible. I am not convinced, however. In Chapter 2 I surveyed and critically analysed three empirical studies that are often cited as evidential support for the illusionist claims that free will scepticism will lead to unethical or antisocial behaviour. In that chapter I explicated salient methodological flaws that show that the assumptions, methods, inferences and conclusions of these studies should be considered unreliable at best. I conclude that to date there is insufficient empirical evidence to support free will illusionism. In the disillusionism camp are the desert sceptics who maintain that belief in free will has a dark side that we would be better off without. The problem they say, is the close connection between believing in free will and the idea that people justly deserve what they get. So rather than having the instrumental benefits that illusionists claim, belief in free will is often called on to justify social and economic inequalities, the harsh treatment of criminals, and other harmful desert based practices. Recent empirical findings from moral and political psychology, as well the evidence presented in this thesis lend strong support to these claims. People who believe more strongly in free will are also more likely to hold right wing authoritarian attitudes, believe in a just world, be more punitive and use moralistic standards for judging self and others.

On balance, there are strong reasons to think that whatever the benefits of retaining lay belief in free will may afford us, the benefits that are likely to come from rejecting this belief are likely to be far greater. Rejecting free will and our desert-based moral responsibility system of retributive justice means rejecting a system that prevents us from looking more deeply into the causes of individual behaviour, that creates and maintains social injustice and poverty, and condones the infliction of enormous suffering, humiliation, and psychological damage on people who do not deserve such treatment. In light is all this, it seems incumbent upon those who recommend persevering the common

view belief in free will to provide a much more convincing case for how perpetuating the belief in free will and desert-based moral responsibility would yield better results than exposing the public to arguments undermining them. As it stands, the extant findings show free will scepticism to be an important and more instrumentally beneficial worldview than believing in free will. I leave you with the words of Caruso who speaks from the heart of free will scepticism (2018a, p. 20):

It's my proposal that we do away with the pernicious belief in free will—and with it the myth of just deserts ... The concepts of free will and desert-based moral responsibility are intimately connected with a number of other potentially harmful beliefs. It's time that we leave these antiquated notions behind, lose our moral anger, stop blaming the victim, and turn our attention to the difficult task of addressing the causes that lead to criminality, poverty, wealth-inequality, and educational inequity.

## BIBLIOGRAPHY

- Appelbaum, L. D. (2002). Who deserves help? students' opinions about the deservingness of different groups living in Germany to receive aid. *Social Justice Research, 15*(3), 201-225.  
doi:<http://dx.doi.org/10.1023/A:1021025528408>
- Adorno, T.W., Frenkel-Brunswik, E., Levinson, D. J., & Sanford, R. N.(1950). *The authoritarian personality*. New York: Harper & Row.
- Allen, J. T., Dimock, M (2007). Pew Research Center for the People & the Press.  
<http://www.pewresearch.org/2007/09/13/a-nation-of-haves-and-havenots/>.
- Altemeyer, B. (1973). *Right-wing authoritarianism*. Winnipeg, Canada: University of Winnipeg Press.
- Bamfield, L., & Horton, T. ( 2009). *Understanding attitudes to tackling economic inequality* . Retrieved from the Joseph Rowntree Foundation  
website: <http://www.jrf.org.uk/sites/files/jrf/attitudes-tackling-economic-inequality-full.pdf> [Google Scholar](#)
- Bargh, J. A. (2008). Free will is un-natural. In J. Baer, J. C. Kaufman & R. F. Baumeister (Eds.), *Are we free? psychology and free will; are we free?*

- psychology and free will* (pp. 128-154, Chapter viii, 356 Pages) Oxford University Press, New York, NY.
- Bargh, J. A., & Ferguson, M. J. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin*, 126(6), 925-945.  
doi:<http://dx.doi.org.helicon.vuw.ac.nz/10.1037/0033-2909.126.6.925>
- Baumeister, R., Masicampo, E., & Dewart, C. (2009). Prosocial Benefits of Feeling Free: Disbelief in Free Will Increases Aggression and Reduces Helpfulness. *Personality and Social Psychology Bulletin*, 35(2), 260-268.
- Baumeister, Roy F., & Brewer, Lauren E. (2012). Believing versus Disbelieving in Free Will: Correlates and Consequences.(Report). *Social and Personality Psychology Compass*, 6, 736.
- Baumeister, Roy F., & Monroe, Andrew E. (2014). Chapter One - Recent Research on Free Will: Conceptualizations, Beliefs, and Processes. In *Advances in Experimental Social Psychology* (Vol. 50, pp. 1-52).
- Black, J. E. (2016). An introduction to the moral agency scale: Individual differences in moral agency and their relationship to related moral constructs, free will, and blame attribution. *Social Psychology*, 47(6), 295-310.  
<http://dx.doi.org.helicon.vuw.ac.nz/10.1027/1864-9335/a000284>
- Brass, M., Lynn, M. T., Demanet, J., & Rigoni, D. (2013). Imaging volition: What the brain can tell us about the will. *Experimental Brain Research*, 229(3), 301-312.  
doi:<http://dx.doi.org.helicon.vuw.ac.nz/10.1007/s00221-013-3472-x>
- Byrne, B. (2010). *Structural equation modeling with AMOS : Basic concepts, applications, and programming* / Barbara M. Byrne. (2nd ed., Multivariate applications book series). New York: Routledge.
- Carey, J. M., & Paulhus, D. L. (2013). Worldview implications of believing in free will and/or determinism: Politics, morality, and punitiveness. *Journal of Personality*, 81(2), 130-141. <http://dx.doi.org/10.1111/j.1467-6494.2012.00799.x>
- Carey, B., and M. Roston. 2015. Three popular psychology studies that didn't hold up. The New York Times. Retrieved from  
<https://www.nytimes.com/interactive/2015/08/28/science/psychology-studies-redid.html>

- Caruso, G. (2013). *Exploring the illusion of free will and moral responsibility* / Gregg D. Caruso. Plymouth, UK: Lexington Books.
- Caruso, G. (2016). Free Will Skepticism and the Question of Creativity: Creativity, Desert, and Self-Creation. *Ergo, an Open Access Journal of Philosophy*, 3(20171122), Ergo, an Open Access Journal of Philosophy, 2016, Vol.3(20171122).
- Caruso, G. D. (2018). Justice without Retribution: An Epistemic Argument against Retributive Criminal Punishment. *Neuroethics*, 1-16.
- Caruso, G. (2018a). Forthcoming in Free Will Skepticism in Law and Society, eds. Elizabeth Shaw, Derk Pereboom, and Gregg D. Caruso. Cambridge University Press.
- Caruso, Gregg. (2018b) "Skepticism About Moral Responsibility", *The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/spr2018/entries/skepticism-moral-responsibility/>>.
- Caruso, G., & Morris, D. (2017). Compatibilism and Retributivist Desert Moral Responsibility: On What is of Central Philosophical and Practical Importance. *Erkenntnis*, 82(4), 837-855.
- Cave, S. (2016, June Issue). There's no such thing as free will: But we're better off believing in it anyway. *The Atlantic*. Retrieved September 30, 2016, from <https://www.theatlantic.com/magazine/archive/2016/06/theres-no-such-thing-as-free-will/480750/>
- Clark, Baumeister, & Ditto. (2017). Making punishment palatable: Belief in free will alleviates punitive distress. *Consciousness and Cognition*, 51, 193-211.
- Clark, Thomas. 2013. Experience and autonomy: Why consciousness does and doesn't matter. In *Exploring the illusion of free will and moral responsibility*, ed. Gregg D. Caruso, 239-254. Lanham, MD: Lexington Books.
- Clark, T (2007). The Scandal of Compatibilism. <https://naturalism.org/resources/book-reviews/the-scandal-of-compatibilism> Retrieved December 18th 2018.
- Crescioni, A., Baumeister, R., Ainsworth, S., Ent, M., & Lambert, N. (2016). Subjective correlates and consequences of belief in free will. *Philosophical Psychology*, 29(1), 1-23.

- Crick, F. (1995). *The astonishing hypothesis : The scientific search for the soul / Francis Crick*. New York, N.Y.: Simon & Schuster.
- Dalbert, C. (1999). *Personal belief in a just world scale*<http://dx.doi.org/helicon.vuw.ac.nz/10.1037/t33337-000> Retrieved from <https://search-proquest-com.helicon.vuw.ac.nz/docview/1948282745?accountid=14782>
- Dalbert, C., Montada, L., & Schmitt, M. (1987). *General belief in just world scale*<http://dx.doi.org/helicon.vuw.ac.nz/10.1037/t06485-000> Retrieved from <https://search-proquest-com.helicon.vuw.ac.nz/docview/1038211896?accountid=14782>
- Deery, O., Davis, T., & Carey, J. (2014). The Free-Will Intuitions Scale and the question of natural compatibilism. *Philosophical Psychology*, 28(6), 1-26.
- Diener, E. (1985). *Satisfaction with life scale*<http://dx.doi.org/helicon.vuw.ac.nz/10.1037/t01069-000> Retrieved from <https://search-proquest-com.helicon.vuw.ac.nz/docview/1038215030?accountid=14782>
- Diener, E., Emmons, R., Larsen, R., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, 49(1), 71-75.
- Eagleman, D. (2011). *Incognito: the secret lives of the brain*. Melbourne, Victoria: The Text Publishing company.
- Earp, B. D., Everett, J. A. C., Nadelhoffer, T., Caruso, G. D., Shariff, A., & Sinnott-Armstrong, W. (2018, April 25). Determined to Be Humble? Exploring the Relationship Between Belief in Free Will and Humility. <https://doi.org/10.31234/osf.io/3bxra>
- Field, A (2011). <https://www.methodspace.com/bonferroni-correcting-lots-of-correlations/>
- Farwell, L., & Weiner, B. (2000). Bleeding hearts and the heartless: Popular perceptions of liberal and conservative ideologies. *Personality and Social Psychology Bulletin*, 26(7), 845-852.  
doi:<http://dx.doi.org/helicon.vuw.ac.nz/10.1177/0146167200269009>

- Genschow, O., Rigoni, D., & Brass, M. (2017). Belief in free will affects causal attributions when judging others' behavior. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 114(38), 10071-10076. <http://dx.doi.org.helicon.vuw.ac.nz/10.1073/pnas.1701916114>
- Gerber, M. M., & Jackson, J. (2013). *Retribution measure*<http://dx.doi.org.helicon.vuw.ac.nz/10.1037/t42617-000> Retrieved from <https://search-proquest-com.helicon.vuw.ac.nz/docview/1721929111?accountid=14782>
- Gerber, M. M., & Jackson, J. (2013). *Treatment of criminal offenders scales*<http://dx.doi.org.helicon.vuw.ac.nz/10.1037/t42618-000> Retrieved from <https://search-proquest-com.helicon.vuw.ac.nz/docview/1721929116?accountid=14782>
- Greene, J., & Cohen, J. (2004). For the law, neuroscience changes nothing and everything. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359(1451), 1775-1785.
- Greene, J., & Cohen, J. (2011). For The Law, Neuroscience Changes Nothing And Everything. *Oxford Handbook of Neuroethics*, Oxford Handbook of Neuroethics.
- Honderich, T. (1988). *A theory of determinism : The mind, neuroscience, and life-hopes / Ted Honderich*. Oxford [England] : New York: Clarendon Press ; Oxford University Press.
- Jost, J. T. (2006). The end of the end of ideology. *American Psychologist*, 61(7), 651-670. <http://dx.doi.org.helicon.vuw.ac.nz/10.1037/0003-066X.61.7.651>
- Kant, I., [1785] & Ellington, J. (1993). *Grounding for the Metaphysics of Morals : With on a Supposed Right to Lie Because of Philanthropic Concerns / Immanuel Kant ; Translated by James W. Ellington*.
- Lambert, A. J., Burroughs, T., & Nguyen, T. (1999). *Belief in a just world scale*<http://dx.doi.org/10.1037/t15579-000> Retrieved from <https://search.proquest.com/docview/1285128610?accountid=14782>

- Libet, B., Gleason, C. A., Wright, E. W, and Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness potential): The unconscious initiation of a freely voluntary act. *Brain* 106: 623-42.
- Lerner, M., & Springer Link. (1980). *The Belief in a Just World A Fundamental Delusion / by Melvin J. Lerner.*, Springer eBooks.
- Levy, N. (2011). Hard luck: How luck undermines free will and moral responsibility. New York: Oxford University Press.
- McCullough, M. E., Emmons, R. A., & Tsang, J. (2002). *Gratitude questionnaire-* <http://dx.doi.org/helicon.vuw.ac.nz/10.1037/t01837-000> Retrieved from <https://search-proquest-com.helicon.vuw.ac.nz/docview/1038210994?accountid=14782>
- McDonald, J.H. 2014. Handbook of Biological Statistics (3rd ed.). Sparky House Publishing, Baltimore, Maryland. This web page contains the content of pages 254-260 in the printed version.
- McGrath, R. E., & Meyer, G. J. (2006). When effect sizes disagree: The case of *r* and *d*. *Psychological Methods*, 11(4), 386-401.  
[doi:http://dx.doi.org/helicon.vuw.ac.nz/10.1037/1082-989X.11.4.386](http://dx.doi.org/helicon.vuw.ac.nz/10.1037/1082-989X.11.4.386)
- Mercier, B. G., Wiwad, D., Piff, P., aknin, L., Robinson, A. R., & Shariff, A. (2018, November 2). Does Belief in Free Will Increase Support for Economic Inequality?. <https://doi.org/10.31234/osf.io/k45ud>
- Miles, J. (2013). 'Irresponsible and a Disservice': The integrity of social psychology turns on the free will dilemma. *British Journal of Social Psychology*, 52(2), 205-218.
- Miles, J. B (2015). The Free Will Delusion. Leicestershire, UK, Matador.
- Monroe, A. E., Dillon, K. D., & Malle, B. F. (2014). Bringing free will down to Earth: People's psychological concept of free will and its role in moral judgment. *Consciousness and Cognition*, 27(1), 100-108.
- Monroe, A. E., Brady, G. L., & Malle, B. F. (2017). This isn't the free will worth looking for: General free will beliefs do not influence moral judgments, agent-specific choice ascriptions do. *Social Psychological and Personality Science*, 8(2), 191-199. <http://dx.doi.org/helicon.vuw.ac.nz/10.1177/1948550616667616>



- Moore, M. (2010). *Placing Blame: A Theory of the Criminal Law*. Oxford: Oxford University Press.
- Morris, S. (2018). The implications of rejecting free will: An empirical analysis. *Philosophical Psychology*, 31(2), 299-321.
- Mueller, C., Dweck, C., & Kruglanski, Arie W. (1998). Praise for Intelligence Can Undermine Children's Motivation and Performance. *Journal of Personality and Social Psychology*, 75(1), 33-52.
- Murray, D., & Nahmias, E. (2014). Explaining Away Incompatibilist Intuitions. *Philosophy and Phenomenological Research*, 88(2), 434-467.
- Nadelhoffer, T. (2011). The Threat of Shrinking Agency and Free Will Disillusionism. In *Conscious Will and Responsibility: A Tribute to Benjamin Libet* (p. ). Oxford University Press.
- Nadelhoffer, T., Shepard, J., Nahmias, E., Sripada, C., & Ross, L. T. (2014). *Free will inventory* <http://dx.doi.org/10.1037/t32432-000> Retrieved from <https://search.proquest.com/docview/1552804325?accountid=14782>
- Nadelhoffer, T., Tocchetto, D. (2013). The potential dark side of believing in free will. *Exploring the Illusion of Free Will and Moral Responsibility*: Lanham, MD: Lexington Books.
- Nadelhoffer, T., & Wright, J. (2018). Humility, Free Will Beliefs, and Existential Angst: How We Got from a Preliminary Investigation to a Cautionary Tale. In *Neuroexistentialism* (p. Neuroexistentialism, Chapter 15). Oxford University Press.
- Nahmias, E., Morris, S., Nadelhoffer, T., & Turner, J. (2005). Surveying Freedom: Folk Intuitions about free will and moral responsibility. *Philosophical Psychology*, 18(5), 561-584.
- Nahmias, E., Morris, S., Nadelhoffer, T., & Turner, J. (2006). Is Incompatibilism Intuitive? *Philosophy and Phenomenological Research*, 73(1), 28-53.
- Nahmias, E (2014). Response to Misirlisoy and Haggard and to Bjornsson and Pereboom. In Sinnott-Armstrong, W. (2014). *Moral Psychology: Free Will and Moral Responsibility*. Cambridge: MIT Press.

- Nahmias, Shepard, & Reuter. (2014). It's OK if 'my brain made me do it': People's intuitions about free will and neuroscientific prediction. *Cognition*, 133(2), 502-516.
- Nichols, Shaun. (2012). The Indeterminist intuition: source and status." *The Monist*, vol. 95, no. 2, pp. 290–307., doi:10.5840/monist201295216.
- Nichols, S., & Knobe, J. (2007). Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions. *Noûs*, 41(4), 663-685.
- Norenzayan, A., & Hansen, I. G. (2006). Belief in supernatural agents in the face of death. *Personality and Social Psychology Bulletin*, 32(2), 174-187.  
<http://dx.doi.org/helicon.vuw.ac.nz/10.1177/0146167205280251> Retrieved from  
<https://search-proquest-com.helicon.vuw.ac.nz/docview/621061392?accountid=14782>
- Oerton, R. T (2018). *The Cruelty of Free Will: How sophistry and savagery supports a false belief*. Leicester, UK, Troubador LTD.
- Overbye, Dennis. (2007). Free will: Now you have it, now you don't. *The New York Times*, January 2.
- Paulhus, D. L., & Carey, J. M. (2011). The FAD-plus: Measuring lay beliefs regarding free will and related constructs. *Journal of Personality Assessment*, 93(1), 96-104. doi:<http://dx.doi.org/10.1080/00223891.2010.528483>
- Paulhus, D. L. and Margesson, A. 1994. *Free Will and Scientific Determinism (FAD-4) scale*, BC, Canada: Unpublished instrument, University of British Columbia, Vancouver.
- Pereboom, D., & Caruso, G. (2018). Hard-Incompatibilist Existentialism: Neuroscience, Punishment, and Meaning in Life. In *Neuroexistentialism* (p. Neuroexistentialism, Chapter 11). Oxford University Press.
- Pereboom, Derk. (2001) *Living Without Free Will*. New York: Cambridge University Press.
- Pereboom, Derk. (2014). *Free Will, Agency, and Meaning in Life*. Oxford Scholarship Online.
- Pommier, E. A., & Neff, K. D. (2010). *Compassion scale*<http://dx.doi.org/helicon.vuw.ac.nz/10.1037/t10177-000> Retrieved from

- <https://search-proquest-com.helicon.vuw.ac.nz/docview/1038217944?accountid=14782>
- Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). *Social dominance orientation scale*<http://dx.doi.org/10.1037/t01146-000> Retrieved from <https://search.proquest.com/docview/1038211096?accountid=14782>
- Rakos, R., Laurene, K., Skala, S., & Slane, S. (2008). Belief in free will: measurement and conceptualization innovations. *Behavior and Social Issues*, 17(1), 20-39.
- Sapp, S. G., & Harrod, W. J. (1993). *Brief locus of control scale*  
doi:<http://dx.doi.org/10.1037/t09281-000>
- Shariff, A., Greene, J., Karremans, J., Luguri, J., Clark, C., Schooler, J., . . . Vohs, K. (2014). Free Will and Punishment: A Mechanistic View of Human Nature Reduces Retribution. *Psychological Science*, 25(8), 1563-1570.
- Shariff, A. F., Schooler, J., & Vohs, K. D. (2008). The hazards of claiming to have solved the hard problem of free will. In J. Baer, J. C. Kaufman, & R. F. Baumeister (Eds.), *Are we free? Psychology and free will* (pp. 181–204). New York: Oxford University Press.
- Sibley, C. G., & Duckitt, J. (2008). Personality and prejudice: A meta-analysis and theoretical review. *Personality and Social Psychology Review*, 12(3), 248-279.  
<http://dx.doi.org/10.1177/1088868308319226>
- Skitka, L. J., & Tetlock, P. E. (1993). Providing public assistance: Cognitive and motivational processes underlying liberal and conservative policy preferences. *Journal of Personality and Social Psychology*, 65(6), 1205-1223.  
doi:<http://dx.doi.org.helicon.vuw.ac.nz/10.1037/0022-3514.65.6.1205>
- Slattery, Trick. (2014). *Breaking the free will illusion; for the betterment of human kind*. Published by Working Matter.
- Smilansky, S. (2000). *Free will and illusion*. New York: Oxford University Press.
- Smilansky, S. (2001). Free Will: From Nature to Illusion. *Proceedings of the Aristotelian Society (Hardback)*, 101(1), 71-95.
- Smith, K. B. (1985). Seeing justice in poverty: The belief in a just world and ideas about inequalities. *Sociol. Spectr.* 5: 17–29.
- Smith, S. M., Roster, C. A., Albaum, G. S., & Golden, L. L. (2016). A multi-group analysis of online survey respondent data quality: Comparing a regular USA

- consumer panel to MTurk samples. *Journal of Business Research*, 69(8), 3139-3148.
- Stillman, T. F., & Baumeister, R. F. (2010). Guilty, free, and wise: Determinism diminish learning from negative emotions. *Journal of Experimental Social Psychology*, 46, 951–960.
- Stillman, T. F., Baumeister, R. F., Vohs, K. D., Lambert, N. M., Fincham, F. D., & Brewer, (2010) Personal philosophy and personnel achievement: Belief in free will predicts better job performance. *Social Psychological and Personality Science*, 1, 43–50.  
doi:<http://dx.doi.org/10.1177/1948550609351600>
- Strawson, G. (1994). The impossibility of moral responsibility. *Philosophical Studies*, 75(1-2), 5–24
- Strawson, Galen. 2010. *Freedom and belief*. Rev. ed. Oxford: Oxford University Press.
- Stroessner, S. J., & Green, C. W. (1989). Effects of belief in free will or determinism on attitudes toward punishment and locus of control. *The Journal of Social Psychology*, 130(6), 789-799.  
doi:<http://dx.doi.org/10.1080/00224545.1990.9924631>
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53-55.
- van Inwagen, P. (1983). *An essay on free will*. Clarendon Press.
- Vargas, M. (2007). Revisionism. In J. Fischer, R. Kane, D. Pereboom, & M. Vargas (Eds.), *Four views on free will* (pp. 126–165). Hoboken, NJ: Wiley.
- Vargas, M. (2013). Situationism and Moral Responsibility: Free Will in Fragments. In *Decomposing the Will* (p. Decomposing the Will, Chapter 17). Oxford University Press.
- Vargas, M. (2015). Desert, responsibility, and justification: A reply to Doris, McGeer, and Robinson. *Philosophical Studies*, 172(10), 2659-2678.
- Viney, W., Waldman, D. A., & Barchilon, J. (1982). Attitudes toward punishment in relation to beliefs in free will and determinism. *Human Relations*, 35(11), 939-950. doi:<http://dx.doi.org/10.1177/001872678203501101>

- Vohs, K. D., & Schooler, J. W. (2008). The value of believing in free will: Encouraging a belief in determinism increases cheating. *Psychological Science*, 19(1), 49-54.  
doi:<http://dx.doi.org/helicon.vuw.ac.nz/10.1111/j.1467-9280.2008.02045.x>
- Waller, B., & ProQuest. (2011). *Against moral responsibility / Bruce N. Waller*. Cambridge, Mass.: MIT Press.
- Waller, B. (2017). Beyond Moral Responsibility to a System that Works. *Neuroethics*, 1-8.
- Waller, Bruce N. (2015). *Stubborn System of Moral Responsibility*. The MIT Press.
- Waller, Bruce. (2013). The Stubborn illusion of moral responsibility. *Exploring the illusion of free will and moral responsibility / Gregg D. Caruso*. Plymouth, UK: Lexington Books. 65-86.
- Walmsley, R. (November 21, 2013). World Prison Population List (tenth edition). International Centre for Prison Studies. Retrieved January 6<sup>th</sup> 2019.
- Ward, K., Longaker, A. J., Williams, J, Naylor., A, Rose, C. H., Simpson, S. G. (2012). Incarceration Within the American and Nordic Prisons: Comparison of National and International Policies.  
<http://www.dropoutprevention.org/engage-backup/incarceration-within-american-and-nordic-prisons/>
- Wegner, D. (2002). *The Illusion of Conscious Will*. Cambridge, MA: Bradford Books, MIT Press.
- Wilson, T. D. (2002). *Strangers to ourselves: Discovering the adaptive unconscious* Belknap Press/Harvard University Press, Cambridge, MA.  
Retrieved from <https://search-proquest-com.helicon.vuw.ac.nz/docview/619950555?accountid=14782>
- Zakrisson, I. (2005). Construction of a short version of the right-wing authoritarianism (RWA) scale. *Personality and Individual Differences*, 39(5), 863-872.  
<http://dx.doi.org/10.1016/j.paid.2005.02.026> Retrieved from  
<https://search.proquest.com/docview/620971687?accountid=14782>
- Zwaan, R. (2013). The value of believing in free will: a replication attempt. March 18, 2013. Retrieved from: <https://rolfzwaan.blogspot.com/2013/03/the-value-of-believing-in-free-will.html>

