

Pace University

DigitalCommons@Pace

---

Pace Law Faculty Publications

School of Law

---

2011

## Punishing Without Free Will

Luis E. Chiesa  
*Pace Law School*

Follow this and additional works at: <https://digitalcommons.pace.edu/lawfaculty>



Part of the [Criminal Law Commons](#), and the [Law and Society Commons](#)

---

### Recommended Citation

Luis E. Chiesa, Punishing Without Free Will, 2011 Utah L. Rev. 1403

This Article is brought to you for free and open access by the School of Law at DigitalCommons@Pace. It has been accepted for inclusion in Pace Law Faculty Publications by an authorized administrator of DigitalCommons@Pace. For more information, please contact [dheller2@law.pace.edu](mailto:dheller2@law.pace.edu).

## PUNISHING WITHOUT FREE WILL

Luis E. Chiesa\*

### I. INTRODUCTION

*Candide: Do you believe that mankind have always been cutting one another's throats; that they were always liars, knaves, treacherous and ungrateful; always thieves, sharpers, highwaymen, lazy, envious and gluttons; always drunkards, misers, ambitious and blood-thirsty; always backbiters, debauchees, fanatics, hypocrites and fools?*

*Martin: Do you not believe that hawks have always preyed upon pigeons, when they could light upon them?*

*Candide: Certainly.*

*Martin: Well, then, if the hawks have always had the same nature, what reason can you give why mankind should have changed theirs?*

*Candide: Oh! There is a great deal of difference; because free will . . . .*<sup>1</sup>

– Voltaire, *Candide*

In a few memorable lines of dialogue between Candide and his faithful servant Martin, Voltaire eloquently conveyed the relevance of free will to our practices of punishing and blaming. In a naturalistic world, as Martin observes, human beings are merely another cog in the wheel of nature. As such, it would be as absurd to blame humans for their sins as it would be to blame hawks for eating pigeons. The always-optimistic Candide begs to differ. The hawk's nature is fixed because animal conduct is determined by natural laws. Mankind's nature, on the other hand, is variable because humans have the ability to change their ways. They have, in other words, free will. And it is because of this uniquely human capacity to freely choose to do otherwise that humans can and should be blamed for their crimes.

Modern criminal law seems to reflect Candide's view of human nature. Free will is central to retributive theories of punishment.<sup>2</sup> For retributivists, the

---

\* © 2011 Luis E. Chiesa. Associate Professor of Law, Pace Law School.

<sup>1</sup> VOLTAIRE, *CANDIDE* ch. XXI (1759), available at <http://www.ourcivilisation.com/smartboard/shop/voltaire/candide/chap21.htm>.

<sup>2</sup> See, e.g., CHRISTOPHER SLOBOGIN, *MINDING JUSTICE: LAWS THAT DEPRIVE PEOPLE WITH MENTAL DISABILITY OF LIFE AND LIBERTY* 9 (2006) (observing that “[t]he so-called free will postulate is an essential premise of the [retributive] model”).

imposition of punishment is justified solely by reference to the offender's deserts.<sup>3</sup> Under the standard view, the offender deserves punishment only if he could have abstained from committing the crime. On the other hand, the offender does not deserve to be punished if he could not have acted otherwise. William Blackstone asserted this view in his famous *Commentaries*, "punishments are . . . only inflicted for the abuse of that free will which God has given to man."<sup>4</sup> In contrast, it is "highly just and equitable that a man should be excused for those acts which are done through unavoidable force and compulsion."<sup>5</sup> This conception of the criminal law continues to influence courts and commentators to this day. As an appellate court explained several decades ago, "our whole criminal code presupposes that an individual possesses a free will and is accountable for his rational conduct."<sup>6</sup>

But what if Martin's view of human nature is more accurate than *Candide's*? What if mankind is no more to blame for their crimes than birds are to blame for eating their prey? As counterintuitive as Martin's deterministic account of human nature might initially sound, many contemporary scientists and philosophers believe it to be closer to the truth than an account of human nature that presupposes the existence of free will.<sup>7</sup> Recent neuroscientific experiments<sup>8</sup> coupled with advances in genetics and related fields<sup>9</sup> increasingly suggest that humans have little control over a wide array of acts that most people believe are freely willed. This has led some respected scholars to contend, as psychologist Daniel Wegner famously noted, that free will is nothing more than an "illusion."<sup>10</sup> Some find these views troubling, as they would call into question our practices of blaming and praising.<sup>11</sup> After all, if every human act is determined by forces that human beings cannot control, how can they be blamed or praised for doing what

---

<sup>3</sup> See generally LEO ZAIBERT, PUNISHMENT AND RETRIBUTION 190 (2006) (discussing the justification of punishment).

<sup>4</sup> IV WILLIAM BLACKSTONE, COMMENTARIES ON THE LAWS OF ENGLAND 1445 (William Draper Lewis ed., 1902).

<sup>5</sup> *Id.*

<sup>6</sup> *State v. Jones*, 577 P.2d 357, 361 (Kan. Ct. App. 1978). See also, more recently, the judgment of the House of Lords in *R v. Kennedy*, [2007] UKHL 38, [14], stating that "[t]he criminal law generally assumes the existence of free will" and that "generally speaking, informed adults of sound mind are treated as autonomous beings able to make their own decisions how they will act."

<sup>7</sup> See, e.g., Galen Strawson, *The Bounds of Freedom*, in THE OXFORD HANDBOOK OF FREE WILL 441 (Robert Kane ed., 2002).

<sup>8</sup> Benjamin Libet et al., *Time of Conscious Intention to Act in Relation to Onset of Cerebral Activity (Readiness Potential): The Unconscious Initiation of a Freely Voluntary Act*, 106 BRAIN, no. 3, 1983 at 623–42.

<sup>9</sup> Regarding the way in which genetics and the environment affect our behavior, see Steven Pinker, *The Blank Slate*, 41 GEN. PSYCHOLOGIST, no. 1, 2006 at 3–4.

<sup>10</sup> DANIEL M. WEGNER, THE ILLUSION OF CONSCIOUS WILL 1 (2002).

<sup>11</sup> Michael S. Moore, *Causation and the Excuses*, 73 CALIF. L. REV. 1091, 1144–45 (1985).

they do?<sup>12</sup> This would appear to cast a mortal blow to our criminal law, for most of the foundational doctrines that undergird our system of criminal justice seem to lose their coherence when unmoored from free will.<sup>13</sup>

Nevertheless, we are faced with mounting scientific evidence against the existence of a uniquely human ability to escape the causal laws of the universe by way of uncaused conscious processes that mysteriously<sup>14</sup> cause human action. Thus, many philosophers now argue that free will is compatible with determinism.<sup>15</sup> They suggest that humans can be free even if it turns out that all events that take place in the universe, including human action, are determined by the interplay between what happened in the past and the present operation of natural laws.<sup>16</sup> According to this account, Martin and Candide's views of human nature are not incompatible. Humans who engage in criminal behavior may possess the sort of free will that makes them responsible for their actions even if their crimes, like the hawk's eating of the pigeon, can be fully explained by the causal laws of nature. But why have philosophers gone out of their way to attempt to reconcile free will with causal determinism? The answer is simple. The truth of determinism seems to threaten the free will edifice upon which our formal and informal institutions of blaming and punishing are built. By demonstrating that free will is compatible with a universe wholly determined by natural laws, the philosopher can rescue the criminal law from the threat of determinism without needing to adopt the unpalatable position of denying determinism's scientific plausibility.

Those who attempt to reconcile free will with determinism appear to believe that many of the things intrinsically important to us, including our practices of blaming, praising, and punishing, would become meaningless in a world without free will.<sup>17</sup> For example, P. F. Strawson believes that certain attitudes central to the human experience such as resentment, gratitude, anger, and love cannot be experienced unless we assume that human beings possess free will.<sup>18</sup> If we are to continue exhibiting the sort of attitudes that make our lives meaningful, we must presuppose that others are responsible for their actions even if scientists

---

<sup>12</sup> *Id.*

<sup>13</sup> *Id.* at 1139, 1144–45.

<sup>14</sup> It is “mysterious” to contend that uncaused conscious processes cause human action because there is presently no scientific account that can explain how this can actually be the case. So far, experimental evidence suggests that human conduct, including conscious processes, is determined by causal events that lie beyond our control.

<sup>15</sup> These philosophers are dubbed “compatibilists.” See, e.g., Christopher Taylor & Daniel Dennett, *Who's Afraid of Determinism? Rethinking Causes and Possibilities*, in THE OXFORD HANDBOOK OF FREE WILL 257.

<sup>16</sup> This is the textbook definition of determinism. See, e.g., Strawson, *supra* note 7, at 442–60.

<sup>17</sup> Moore, *supra* note 11, at 1144–45.

<sup>18</sup> Peter Strawson, *Freedom and Resentment*, 48 PROC. BRITISH ACAD. 1, 1–25 (1962), reprinted in PERSPECTIVES ON MORAL RESPONSIBILITY 45, 45–67 (John Martin Fischer & Mark Ravizza eds., 1993).

demonstrate that human action is determined by causal forces beyond human control.<sup>19</sup> Similarly, Michael Moore has argued that free will is essential to us because “[o]ur moral life is built upon our praising or blaming people when they help a friend, tell a bad joke, create a work of art, or write a clear and truthful essay.”<sup>20</sup> As a result, rejecting free will is undesirable because it would lead to “falsify[ing] much of our moral life.”<sup>21</sup>

This Article will argue that there are good moral reasons to conclude that the scientific plausibility of determinism ought to lead us to *abandon* the notion of free will. *Contra* P. F. Strawson and Moore, this Article suggests that rejecting free will does not undermine the human experience, and doing so is plausible and attractive because it would likely lead to more humane and efficient institutions of blaming and punishing. The argument consists of six parts.

Part II fleshes out in more detail the claim that our criminal laws presuppose the existence of freely willed actors who are capable of meaningfully controlling their conduct. This simple postulate seems to lie at the heart of many foundational doctrines of criminal law, including the voluntary act requirement,<sup>22</sup> the insanity defense,<sup>23</sup> and the general theory of excuse defenses.<sup>24</sup> It also undergirds certain important features of the proximate cause<sup>25</sup> and *mens rea*<sup>26</sup> doctrines. It is therefore difficult to deny that free will is embedded in the very fabric of our system of criminal justice.

Part III examines the current state of the philosophical debates on free will. After discussing the literature, this Article concludes that the plausibility of the thesis of causal determinism is undeniable. At the very least, it is unclear whether humans actually have the ability to consciously control their conduct.<sup>27</sup> The plausibility of determinism has triggered diverse reactions from the philosophical

---

<sup>19</sup> *Id.*

<sup>20</sup> Moore, *supra* note 11, at 1144.

<sup>21</sup> *Id.* at 1045.

<sup>22</sup> 1 GEORGE P. FLETCHER, *THE GRAMMAR OF CRIMINAL LAW: AMERICAN, COMPARATIVE, AND INTERNATIONAL* 273–81 (2007).

<sup>23</sup> Under the Model Penal Code formulation of the insanity defense, an actor is not criminally liable if he lacks the capacity to conform his conduct to the mandates of the law. MODEL PENAL CODE § 4.01.

<sup>24</sup> *See, e.g.*, Joshua Dressler, *Rethinking the Heat of Passion: A Defense in Search of a Rationale*, 73 J. CRIM. L. & CRIMINOLOGY 421 (1982).

<sup>25</sup> According to general principles of proximate causation, “[t]he *free*, deliberate, and informed intervention of a second person, who intends to exploit the situation created by the first, but is not acting in concert with him, is normally held to relieve the first actor of criminal responsibility.” H.L.A. HART & TONY HONORÉ, *CAUSATION IN THE LAW* 326 (1985) (emphasis added); *see also* Glanville Williams, *Finis for Novus Actus?*, 48 CAMBRIDGE L.J. 391, 392 (1989).

<sup>26</sup> According to section 2.02 of the Model Penal Code, for example, a person acts “purposely” if his “conscious objective” is to commit the crime. MODEL PENAL CODE § 2.02 (2001).

<sup>27</sup> Manuel Vargas, *Response to Kane, Fischer and Pereboom*, in *FOUR VIEWS ON FREE WILL* 204, 206 (Ernest Sosa ed., 2007).

community. Libertarians believe that determinism is incompatible with free will and that the recent literature, therefore, threatens to undermine free will.<sup>28</sup> They argue, however, that humans are free because their decisions are ultimately grounded in irreducibly indeterministic processes, which are likely explained by the principles undergirding the field of quantum mechanics.<sup>29</sup> Hard incompatibilists, on the other hand, believe that determinism is probably true and that both determinism and indeterminism are incompatible with free will.<sup>30</sup> Finally, compatibilists believe that humans possess free will regardless of the truth of causal determinism.<sup>31</sup> The primary goal of this part is to explore in detail the positions espoused by libertarians, hard incompatibilists, and compatibilists.

Part IV looks at what legal scholars have to say about the free will debate and its relevance to criminal theory. Despite the obvious threat that recent scientific literature on the nature of consciousness poses for the conventional “free will centered” account of criminal law, legal theorists have paid surprisingly little attention to the subject. Most seem content to casually point out that it is unclear whether human beings have the kind of free will that our criminal laws assume we have. They then proceed to nonchalantly dismiss the problem as one that ought to be of concern for philosophers rather than lawyers. This is the tack taken by Douglas Husak in his book chapter on the criminal law’s voluntary act requirement.<sup>32</sup> Others, like George Fletcher, have attempted to defend a libertarian account of free will by arguing that Noam Chomsky’s theory of language lends credence to the proposition that human action is inherently indeterminate.<sup>33</sup> In contrast, some, like Stephen Morse,<sup>34</sup> Kim Ferzan,<sup>35</sup> and Michael Moore,<sup>36</sup> adopt the classic compatibilist stance and argue that the sort of free will that is needed to make sense of the criminal law is not at odds with determinism. Finally, there are those who, like Larry Alexander,<sup>37</sup> grant that determinism is incompatible with the

---

<sup>28</sup> See, e.g., Robert Kane, *Some Neglected Pathways in the Free Will Labyrinth*, in THE OXFORD HANDBOOK OF FREE WILL 406.

<sup>29</sup> Taylor & Dennett, *supra* note 15, at 259.

<sup>30</sup> See, e.g., Derk Pereboom, *Living Without Free Will: The Case for Hard Incompatibilism*, in THE OXFORD HANDBOOK OF FREE WILL 477.

<sup>31</sup> Ishtiyaque Haji, *Compatibilist Views of Freedom and Responsibility*, in THE OXFORD HANDBOOK OF FREE WILL 202.

<sup>32</sup> DOUGLAS N. HUSAK, *PHILOSOPHY OF CRIMINAL LAW* 78–122 (1987). After concluding that the law ought to abandon the voluntary act doctrine in favor of what he dubs the “control requirement,” Husak concedes that a strand of the philosophical literature on free will casts doubt on whether humans have the ability to exert meaningful control over their actions. *Id.* at 98. Surprisingly, however, he goes on to claim that this is a subject that he need not delve into. *Id.*

<sup>33</sup> FLETCHER, *supra* note 22, at 273–81 (2007).

<sup>34</sup> See, e.g., Stephen Morse, *Criminal Responsibility and the Disappearing Person*, 28 CARDOZO L. REV. 2545 (2007).

<sup>35</sup> LARRY ALEXANDER & KIMBERLY KESSLER FERZAN, *CRIME AND CULPABILITY: A THEORY OF CRIMINAL LAW* 15 (2009).

<sup>36</sup> Moore, *supra* note 11.

<sup>37</sup> ALEXANDER & FERZAN, *supra* note 35, at 15 n.22.

structure of criminal law, but mysteriously declare themselves agnostics with regard to the truth of determinism.

Given the intractable nature of the free will problem, Part V argues that the scientific, philosophical, and legal debate on the subject has led to a dialectical stalemate. A dialectical stalemate arises when proponents of an argument cannot muster sufficient evidence to decisively prove the argument, while the argument's detractors cannot gather sufficient evidence to decisively refute it.<sup>38</sup> Those who claim that humans have enough control over their actions to be considered morally responsible for their conduct have failed to conclusively demonstrate either that determinism is probably false<sup>39</sup> or that the truth of determinism does not undermine their claim.<sup>40</sup> In contrast, those who suggest that determinism is probably true and that it is incompatible with the free will necessary for moral responsibility have failed to decisively demonstrate the latter.<sup>41</sup>

How are we to proceed in light of this dialectical stalemate? Some criminal theorists have thrown their hands up in despair and opted to favor the status quo over radically revising the criminal laws to make them compatible with the increasingly likely possibility that determinism is true.<sup>42</sup> In contrast, Part V argues that the best way to move beyond the stalemate is to put forth normative arguments to demonstrate which of the competing solutions to the free will problem makes our life in general—and our practices of blaming and punishing in particular—more appealing. After all, if the way out of the free will labyrinth is unclear, why not embark on the most attractive path until we finally stumble upon evidence that decisively demonstrates that we are headed in the wrong direction?

Part VI contends that a life without free will is not as bad as most philosophers and legal scholars would have us believe. Taking a cue from Derk Pereboom's writings on the subject,<sup>43</sup> this Article suggests that it is not altogether clear that rejecting the notion of free will inevitably leads to renouncing certain attitudes that are inextricably part of the human experience. For example, consider love. Some philosophers have argued that love loses most of its meaning if we assume that others are not free to choose whether to love us and we are not free to determine whether to love them in return.<sup>44</sup> This is far from obvious, as the case of small children illustrates. We truly and meaningfully love our newborn children even if we believe, as we should, that they are not (yet) capable of freely choosing to love us back. It would also seem that we will not stop loving our spouses, friends, and extended family merely because it is demonstrated to us that we did

---

<sup>38</sup> JOHN MARTIN FISCHER, *OUR STORIES: ESSAYS ON LIFE, DEATH, AND FREE WILL* 116–17 (2009).

<sup>39</sup> This is what libertarians like Kane must prove.

<sup>40</sup> This is what compatibilists like Dennett must demonstrate.

<sup>41</sup> This is what hard incompatibilists like Pereboom must prove. In light of recent scientific evidence, it would seem that hard determinists stand on solid footing regarding their claim that determinism is probably true.

<sup>42</sup> ALEXANDER & FERZAN, *supra* note 35, at 15 n.22.

<sup>43</sup> See DERK PEREBOOM, *LIVING WITHOUT FREE WILL* (2001).

<sup>44</sup> Strawson, *supra* note 18, at 53.

not freely choose to feel what we feel for them. The same is true with most of the feelings that make our life worth living.

Part VI also argues that, far from depriving life of meaning, the assumption that humans lack free will may actually make life *more* appealing. First, the practice of blaming other people for their sins and crimes loses meaning in a world without free will. Contrary to what some have argued, living without blaming others for their wrongs would not be fatal to interpersonal relationships. If it is true, as it seems to be, that human action is determined by factors they cannot control, it would intuitively follow that humans should not be blamed for the consequences of their conduct. Furthermore, even if we assume that it is fair to blame humans for engaging in conduct that they cannot control, it is unclear whether doing so is beneficial for mankind. After all, our practices of blaming have historically been linked to unappealing attitudes of resentment, anger, and vengeance. A world that does not encourage the formation of those attitudes would, at the very least, be less violent and cruel towards those who engage in wrongdoing. In light of the increasingly inhumane treatment that criminals receive both in this country and abroad,<sup>45</sup> this would appear to be a welcome development.

In Part VII, this Article argues that rather than leading to the disappearance of criminal law, assuming that humans lack free will might lead to a more appealing system of criminal justice.<sup>46</sup> Once we presuppose that humans do not have enough control over their acts to be considered responsible for their conduct, it is clear that imposing punishment on wrongdoers for the purposes of exacting retribution is unwarranted. As a result, in a world without free will the purpose of punishment must shift from giving to the offender what he deserves, to protecting society from dangerous individuals who are nevertheless not to blame for their transgressions.

This Article therefore contends that punishment should be reconceptualized as a type of quarantine in which individuals who are not responsible for the conditions that make them dangerous are deprived of certain liberties for the protection of others.<sup>47</sup> Furthermore, just as the government has the responsibility to treat a person quarantined with swine flu in a manner that ensures her speedy recovery and encourages her return to life in society, the state should also have the

---

<sup>45</sup> Our current system of criminal justice increasingly relies on incarceration as the preferred method of punishment for many of the most pervasive offenses committed by individuals today, such as weapon and drug possession offenses. This has generated a severe overcrowding problem in America's prisons. This, in turn, leads to the inhumane treatment of inmates in a variety of ways, including, but not limited to, the failure to provide them with adequate medical care, the creation of unsafe and unsanitary living conditions, and the creation of an environment which promotes violence and unrest amongst the prison population. *Brown v. Plata*, 131 S. Ct. 1910, 1923–24 (2011). The magnitude of the problem is such that the United States Supreme Court recently declared that the overcrowding problem in the California prison system has transformed the punishment imposed on these inmates into one that is "cruel and unusual" in violation of the Eighth Amendment. *Id.*

<sup>46</sup> See *infra* notes 335–374 and accompanying text.

<sup>47</sup> See *infra* notes 347–350 and accompanying text.



duty to treat someone quarantined for committing an offense in a way that increases the likelihood that he will be able to return to societal life.<sup>48</sup>

A system of criminal justice that does not assume that humans possess free will would also rely less on incarceration. We send people to jail mostly because we believe that they deserve to suffer for what they have done. It would be cruel, however, to intentionally make people suffer by throwing them in jail for doing something that they could not have abstained from doing. Consequently, if we assume that actors lack free will, it should lead us to seek alternatives to incarceration as a response to crime. This would spur a more efficient criminal law, as the resources that are consumed by our costly prison system could be diverted to more socially useful and humane methods of social control.

Finally, the Article concludes by urging criminal lawyers and theorists to take seriously the idea that we might not have the sort of free will that informs most of our foundational criminal law doctrines. They should do this for two reasons. First, the scientific evidence increasingly suggests that humans lack the capacity to control their acts. Second, a criminal justice system that presupposes that human beings lack free will is more humane and efficient than one that assumes humans are free to act as they please. As a result, there are good reasons to believe that the most attractive way out of the free will maze is to assume that determinism and indeterminism are incompatible with the type of freedom that lies at the core of contemporary criminal law and theory.

## II. THE IMPORTANCE OF FREE WILL TO OUR CURRENT PRACTICES OF BLAMING AND PUNISHING

Philosophers and scientists have long debated whether humans possess the sort of free will that undergirds *Candide's* account of human nature.<sup>49</sup> Although the debate rages in academic circles, most people assume that we do have the ability to freely choose our actions, and that, as a result of this capacity, it makes sense to blame or praise us for engaging in certain acts. This assumption is not only prevalent, but also deeply held—belief in free will is central to many of the practices that define our society. Free will is an essential tenet of Judeo-Christian morality.<sup>50</sup> It is also central to our everyday practice of holding our friends and

---

<sup>48</sup> See *infra* notes 356–361 and accompanying text.

<sup>49</sup> The modern formulation of the free will problem is sometimes traced back to the writings of Epicurus. See, e.g., Epicurus, *Letter to Menoeceus*, in CLASSICS OF MORAL AND POLITICAL THEORY 455 (2d ed. 1992) (“[N]ecessity is not answerable [to anyone], chance is unstable; while what occurs by our own agency is autonomous, and that it is to this that praise and blame are attached.”). Medieval scholars also addressed the free will problem. See, e.g., 1 THOMAS AQUINAS, SUMMA THEOLOGICA, question 83, art. 1 (Fathers of the English Dominican Province trans., 2d rev. ed. 1920). Modern discussions of the problem are legion. For a compilation containing representative modern discussions of the free will problem, see DERK PEREBOOM, FREE WILL (Hackett Pub., 2d ed., 2009).

<sup>50</sup> According to Judeo-Christian thought, God created a world full of alternative courses of action, some of which are good and some of which are bad. Furthermore, he

loved ones accountable for what they do. When we chide our best friend for not keeping a secret, we typically do so as a way of expressing our firmly held conviction that she should have (and *could* have) abstained from acting the way she did. Given that belief in free will pervades so many aspects of our lives, it should come as no surprise that it is also essential to understanding our current approach to criminal law. In the following pages, this Article discusses some ways in which belief in free will shapes our practices of blaming and punishing.

### A. *The Voluntary Act Requirement*

As any first-year law student knows, there can be no criminal liability without proof that the defendant engaged in the *actus reus* of the offense charged.<sup>51</sup> *Actus reus* is the Latin term for the “guilty act” which must take place if an individual is to be subjected to criminal punishment.<sup>52</sup> It is a basic principle of criminal law that an act is “guilty,” and therefore punishable, only if it is *voluntary*.<sup>53</sup> This has come to be known among courts and commentators as the criminal law’s so-called “voluntary act principle or requirement.”<sup>54</sup> Pursuant to this principle, there can be no criminal liability for involuntary acts that are “not a product of the effort or the determination of the actor.”<sup>55</sup> As a result, harm brought about by a bodily movement that is the product of a reflex, convulsion, or sleepwalking does not trigger the imposition of criminal liability.<sup>56</sup> Furthermore, an individual may only be held liable for *what he does* rather than for *who he is*. As a result, the criminal law may legitimately prohibit the act of acquiring cocaine, but not the status of being a drug addict.<sup>57</sup>

The nexus between the voluntary act requirement and free will is evident. An act may lack free will if it is the product of the actor’s volition but internal or

---

endowed humans with the capacity to freely choose amongst these alternative courses of action. *See, e.g., Deuteronomy* 30:19 (“I [God] have set before you life and death, blessing and cursing: therefore choose life . . .”).

<sup>51</sup> *See* HUSAK, *supra* note 33, at 78.

<sup>52</sup> BLACK’S LAW DICTIONARY 41 (9th ed. 2009).

<sup>53</sup> *E.g.*, MODEL PENAL CODE § 2.01(1) (2001).

<sup>54</sup> *See generally* H.L.A. HART, PUNISHMENT AND RESPONSIBILITY 97 (1968) (tracing the legal understanding of voluntary action to John Austin, who theorized that a human action is a “muscular contraction” caused by a “volition” or an “act of will”).

<sup>55</sup> MODEL PENAL CODE § 2.01(2)(d).

<sup>56</sup> According to the MODEL PENAL CODE section 2.01(2), reflexes, convulsions and bodily movements during unconsciousness or sleep “are not voluntary acts within the meaning of this Section.” *See also* *State v. Sowry*, 803 N.E.2d 867, 870 (Ohio Ct. App. 2004) (“[A]cts performed while unconscious or sleepwalking are not voluntary acts [that generate criminal liability.]”); *State v. Case*, 672 A.2d 586, 589 (Me. 1996) (“To be voluntary an act must be the result of an exercise of defendant’s conscious choice to perform [it], and not the result of reflex [or] convulsion . . .”).

<sup>57</sup> The prohibition of status crimes is so fundamental that the Supreme Court has held that the Eighth Amendment proscribes punishing mere status. *Robinson v. California*, 370 U.S. 660, 666–67 (1962).

external pressures undermine the actor's volition significantly.<sup>58</sup> On the other hand, an act may also lack free will when it is not the product of the actor's volition.<sup>59</sup> Conduct that fails to satisfy the criminal law's voluntary act requirement is an example of the latter type of conduct. Involuntary acts are thus not reflective of free will. This is one of the reasons why the voluntary act requirement is such an essential feature of our criminal law. As the Seventh Circuit noted:

In the narrowest sense, every crime must be the product of defendant's free will; it must reflect his choice to perform the criminal act. If the act itself was the result of a mere reflex, or muscular spasm, or was caused by physical duress or compulsion, even the narrowest intent would be absent and the defendant would be innocent of crime; indeed, it could be said that he did not act at all. It is in this sense that the traditional defense of "compulsion" or "necessity" may justify an act that would be unlawful if it had reflected a deliberate exercise of the defendant's free will.<sup>60</sup>

The voluntary act principle reflects the deeply held belief that it is unfair to punish someone for engaging in acts that are not the product of a free will.

### *B. Legal Causation*

If a defendant is charged with a crime that brings about a harmful consequence, she can only be held liable if her conduct *caused* the harmful consequence to materialize. A defendant's conduct is considered a cause of the harm only if it is both an "actual" and "legal" (or "proximate") cause of the harm.<sup>61</sup> Conduct is an "actual" cause of harm if the harm would not have occurred but for the defendant's act.<sup>62</sup> The standard for determining whether the defendant's conduct was a "legal" cause of the harm is not settled. Some courts have held that

---

<sup>58</sup> In his *Nicomachean Ethics*, Aristotle points out that there is a sense in which acts done under extreme pressure, such as throwing cargo overboard during a storm in order to save oneself, are not voluntary and thus not reflective of free will, for no one would choose any such act in itself. ARISTOTLE, *NICOMACHEAN ETHICS* 52–53 (Martin Ostwald trans., Bobbs-Merrill Co. 1962).

<sup>59</sup> According to Robert Nozick, for example, we lack free will if we are not the originators of our acts. ROBERT NOZICK, *PHILOSOPHICAL EXPLANATIONS* 291–92 (1981). Therefore, we lack free will when our acts are the product of something other than our volitions.

<sup>60</sup> *United States v. Cullen*, 454 F.2d 386, 390–91 (7th Cir. 1971).

<sup>61</sup> *See, e.g., State v. Rivas*, 896 P.2d 57, 62 (Wash. 1995) ("In crimes which are defined to require specific conduct resulting in a specified result, the defendant's conduct must be the "legal" or "proximate" cause of the result. Before criminal liability is imposed, the conduct of the defendant must be both (1) the actual cause, and (2) the "legal" or "proximate" cause of the result." (citation omitted)).

<sup>62</sup> *E.g., People v. Zak*, 457 N.W.2d 59, 63 (Mich. Ct. App. 1990) (citing a well-known criminal law treatise, the court stated that "[a] factor is an *actual cause* of a result if the result would not have occurred when it did in the absence of that factor").

determinations of legal causation hinge on whether the harm was foreseeable in light of defendant's conduct,<sup>63</sup> while others have suggested that legal causation depends on whether the harmful result can be directly traced back to the defendant's conduct.<sup>64</sup> In contrast, the Model Penal Code suggests that the defendant's conduct ought to be considered the legal cause of the harm unless the connection between the act and the harm that resulted is too tenuous to have a just bearing on the defendant's liability.<sup>65</sup> Regardless of which view of proximate cause is adopted, courts and commentators agree that an event that takes place after the defendant has acted but before the harm has materialized may sever the link between the defendant's conduct and the harm.<sup>66</sup> There is also agreement that such intervening events are more likely to break the causal connection between the defendant's act and the harm if the events involve a voluntary act by the victim or a third party.<sup>67</sup>

Examples abound that illustrate the relevance of the victim's voluntary conduct to determinations of legal causation. Perhaps the most common example involves an actor who provides another person the means with which to commit suicide.<sup>68</sup> Does this conduct constitute homicide? The general rule in these cases is that the defendant is not liable for homicide because the victim's voluntary decision to commit suicide constitutes an intervening cause, which breaks the causal link between the defendant's action (supplying the means to commit suicide) and the result that ensued (death of the victim).<sup>69</sup> Given that the victim's conduct is irrelevant to the criminal law,<sup>70</sup> why do the victim's voluntary acts determine the defendant's liability in this context? It seems that the voluntary conduct of the victim is important in these cases because freely willed acts occupy a special place in our legal judgments about the causal connections between certain events. For better or worse, the law of causation reflects a deeply held belief that freely willed acts have more causal force than non-freely willed acts. As one Alabama court stated, "[a] determination as to whether the conduct of a person

---

<sup>63</sup> *E.g.*, *Williams v. State*, 235 S.W.3d 742, 764 (Tex. Crim. App. 2007) ("Obviously, some element of foreseeability limits criminal causation just as it limits principles of civil 'proximate causation.'").

<sup>64</sup> *See, e.g.*, *People v. Kibbe*, 321 N.E.2d 773, 773–74 (N.Y. 1974) (asserting that the defendants should not be found guilty unless their conduct "was a cause of death sufficiently direct as to meet the requirements of the *criminal*, and not the *tort*, law").

<sup>65</sup> MODEL PENAL CODE § 2.03(2) (2001).

<sup>66</sup> The reference is, of course, to the doctrine of intervening causation. For a discussion of intervening causation in the criminal law, see *Starkenburg v. State*, 934 P.2d 1018, 1022–26 (Mont. 1997) (discussing intervening acts by a third-party).

<sup>67</sup> *See generally* PAUL H. ROBINSON, CRIMINAL LAW 162 (1997) (discussing cases in which a voluntary intervening act by the victim was found to break the causal connection).

<sup>68</sup> *See, e.g.*, *People v. Kevorkian*, 527 N.W.2d 714, 735–39 (Mich. 1994).

<sup>69</sup> *See id.* at 445.

<sup>70</sup> *See generally* VERA BERGELSON, VICTIMS' RIGHTS AND VICTIMS' WRONGS: COMPARATIVE LIABILITY IN CRIMINAL LAW 9–35 (2009) (explaining that the victim's conduct is irrelevant except for a few exceptions).

caused the suicide of another must necessarily include an examination of the victim's free will," since "[c]ases have consistently held that the 'free will of the victim is seen as an intervening cause which . . . breaks the chain of causation' between the defendant's act and the victim's death."<sup>71</sup>

### C. *Mens Rea*

Edward Coke's often-cited assertion that *actus non facit reum nisi mens sit rea* is as true today as it was when it was first asserted.<sup>72</sup> The Latin maxim means that an "act does not make a person guilty unless their mind is also guilty." This mens rea requirement has generally led to punishing the occurrence of harm only if the actor intended to bring about the harm.<sup>73</sup> Although punishing negligent wrongdoing is not unheard of, most core criminal offenses can only be committed intentionally.<sup>74</sup> Criminalization of negligently caused harm is mostly relegated to so-called public welfare offenses whose commission does not usually stigmatize the offender in the same way as the commission of a core crime such as rape or theft.<sup>75</sup> In spite of the traditional reluctance to punish negligent harm, some core offenses can be committed negligently. The most salient example is the criminalization of negligent homicide, which is considered an offense in every American jurisdiction.<sup>76</sup> However, even when a harm caused by a negligent act is criminalized, it is always punished less severely than an intentional act causing the same kind of harm.<sup>77</sup>

It is difficult to explain criminal law's unwillingness to punish negligent wrongdoing, especially since, as tort law illustrates, damage caused by the negligent wrongdoer is, all things being equal, identical to damage caused by an

---

<sup>71</sup> Lewis v. State, 474 So. 2d 766, 771 (Ala. Crim. App. 1985) (quoting S.W. Brenner, *Undue Influence in the Criminal Law: A Proposed Analysis of the Criminal Offense of "Causing Suicide,"* 147 ALB. L. REV. 62, 83 (1982)).

<sup>72</sup> 3 EDWARD COKE, INSTITUTES ON THE LAWS OF ENGLAND: CONCERNING TREASON AND OTHER PLEAS OF THE CROWN AND CRIMINAL CAUSES, at c.1, fo.10 (1797).

<sup>73</sup> See, e.g., *Morrisette v. United States*, 342 U.S. 246, 250–51 (1952) ("[A]n injury can amount to a crime only when inflicted by intention.").

<sup>74</sup> This is the case, for example, with the offenses of rape, robbery, burglary and kidnapping. See MODEL PENAL CODE § 213.1(1) (2001) (Rape); *id.* § 222.1 (Robbery); *id.* § 221.1 (Burglary); *id.* § 212.1 (Kidnapping). When the definition of an offense does not include a mental state, the offense can only be committed purposely, knowingly, or recklessly, but not negligently. *Id.* § 2.02(3).

<sup>75</sup> The classic discussion of public welfare offenses is still Francis B. Sayre, *Public Welfare Offenses*, 33 COLUM. L. REV. 55 (1933).

<sup>76</sup> The Model Penal Code's negligent homicide provision is illustrative of the typical negligent homicide statute. See MODEL PENAL CODE § 210.4.

<sup>77</sup> Negligent homicide is always punished less severely than intentional homicide. In New York, for example, negligent homicide is a Class E felony, whereas intentional homicide is either a Class B or Class A felony. Compare N.Y. PENAL LAW § 125.10 (McKinney 2009) (criminally negligent homicide), with *id.* §§ 125.25–125.27 (versions of intentional homicide).

intentional wrongdoer.<sup>78</sup> For example, an actor who kills a victim by deliberately poisoning her inflicts the same harm as the actor who kills a victim by carelessly running her over with his car, regardless of the actor's mental state when committing the offense. Nevertheless, the careless killing will be punished much less severely than the intentional killing. The standard explanation for this differential treatment is that an actor who intentionally causes harm is more blameworthy than an actor who negligently causes the same harm.<sup>79</sup> After all, as Oliver Wendell Holmes pointed out when discussing the different moral quality of intentional and negligent acts, "even a dog distinguishes between being stumbled over and being kicked."<sup>80</sup>

The perceived differences in the blameworthiness of these acts are inextricably linked to beliefs about the degree of free will with which intentional and negligent wrongdoers act. Intentional harm causation is thus believed more deserving of condemnation than negligent harm causation because, as the Supreme Court stated in *Morrisette v. United States*, of the criminal law's "universal and persistent . . . belief in freedom of the human will and a consequent ability and duty of the normal individual to choose between good and evil."<sup>81</sup> As a result of this belief, some argue that "our substantive criminal law is based upon a theory of punishing the vicious will," for it assumes that the paradigmatic case of wrongdoing is that of a "free agent confronted with a choice between doing right and doing wrong and freely choosing . . . to do wrong."<sup>82</sup>

#### D. Legal Insanity

In 1843, Daniel M'Naghten attempted to assassinate the British Prime Minister.<sup>83</sup> Given that M'Naghten claimed that he committed the crime while in a state of mental incapacity, the House of Lords came up with a set of rules for determining when a defendant should be acquitted as a result of insanity. The rule set out by the House of Lords in the case, known as the *M'Naghten* test for legal insanity, holds that a defendant should be relieved of criminal liability only if at the time of the crime he suffered from a defect of reason or mental disease that

---

<sup>78</sup> Punitive damages can be imposed on the intentional wrongdoer. This aspect of tort law, as the term "punitive" implies, more closely resembles criminal than civil law. See *Smith v. Wade*, 461 U.S. 30, 48 (1983) (recognizing that "punitive damages in tort cases may be awarded not only for actual intent to injure or evil motive, but also for recklessness, serious indifference to or disregard for the rights of others, or even gross negligence").

<sup>79</sup> See generally JEROME HALL, *GENERAL PRINCIPLES OF CRIMINAL LAW* 135–41 (1960) (evaluating the various theories behind punishing negligent homicide).

<sup>80</sup> OLIVER WENDELL HOLMES, JR., *THE COMMON LAW* 3 (1881).

<sup>81</sup> *Morrisette v. United States*, 342 U.S. 246, 250 (1952).

<sup>82</sup> Roscoe Pound, *Introduction* to FRANCIS B. SAYRE, *A SELECTION OF CASES ON CRIMINAL LAW*, at xxix, xxxvi–xxxvii (1927).

<sup>83</sup> Michael Stoll, Note, *Miles to Go Before We Sleep: Arizona's "Guilty Except Insane" Approach to the Insanity Defense and Its Unrealized Promise*, 97 *GEO. L.J.* 1767, 1772 (2009).

prevented him from appreciating the nature, quality, or wrongfulness of his act.<sup>84</sup> Several decades after M’Naghten’s case was decided, courts started invoking a different standard of insanity that has been dubbed the “irresistible impulse” test.<sup>85</sup> According to the irresistible impulse test, a defendant should be acquitted if at the time of the commission of the offense he suffered from a mental disease or condition that prevented him from controlling his conduct.<sup>86</sup>

Subsequently, the drafters of the Model Penal Code decided to combine the *M’Naghten* and “irresistible impulse” tests.<sup>87</sup> Under the Model Penal Code, a defendant will be acquitted for reason of insanity if at the time of the crime he suffered from a mental condition or defect that caused him to lack substantial capacity<sup>88</sup> either to appreciate the wrongfulness<sup>89</sup> of his conduct *or* to conform his conduct to the mandates of the law. Although the *M’Naghten* test continues to be applied in many American jurisdictions, a substantial number of states apply the Model Penal Code’s combined *M’Naghten-Irresistible Impulse* test.<sup>90</sup>

The *M’Naghten* test of legal insanity illustrates the role of free will in criminal law. The test focuses on whether the cognitive capabilities of the defendant were undermined by mental disease or defect in a way that prevented him from appreciating the import of his conduct and the consequences of his acts. These cognitive impairments preclude the imposition of liability because, as one judge put it, they weaken “the power in man to make a choice between alternative courses of action.”<sup>91</sup> Thus, courts frequently construe the *M’Naghten* rules as a vehicle to determine whether the defendant exercised the sort of free will that undergirds our conventional practices of blaming and punishing. Consequently,

---

<sup>84</sup> M’Naghten’s Case, 8 Eng. Rep. 718, 722–23 (1843).

<sup>85</sup> The irresistible impulse test can be traced back to *Parsons v. State*, 2 So. 854 (Ala. 1887).

<sup>86</sup> *Id.* The test has also come to be known as the “policeman at the elbow” test, for it is asserted that someone commits a crime pursuant to an irresistible impulse if he would have committed the offense even if a police officer was standing next to him. For a relatively recent discussion of the “policeman at the elbow” test, see *People v. Jackson*, 627 N.W.2d 11, 13–14 (Mich. Ct. App. 2001).

<sup>87</sup> MODEL PENAL CODE § 4.01(1) (2001).

<sup>88</sup> The Model Penal Code standard broadens the scope of the insanity defense, for it merely requires a showing of a lack of “substantial capacity” to appreciate the wrongfulness of the conduct or to control the act rather than a showing of a complete absence of knowledge (*M’Naghten*) or a total lack of control over the conduct (irresistible impulse). *See id.*

<sup>89</sup> The drafters of the Model Penal Code pointed out that states drafting an insanity defense have to choose between requiring that the actor “lack capacity to appreciate” the wrongfulness of his conduct *or* demanding that the actor lack capacity to understand the *criminality* of his conduct. *Id.* The difference is significant. An actor does not appreciate the wrongfulness of his conduct if he fails to see that his act is legally or morally wrongful. In contrast, an actor does not appreciate the *criminality* of his conduct if he fails to see that his conduct is against the law.

<sup>90</sup> *See* 2 CHARLES E. TORCIA, WHARTON’S CRIMINAL LAW § 104 (15th ed. 1994).

<sup>91</sup> *State v. Esser*, 115 N.W.2d 505, 529 (Wis. 1962) (Hallows, J., dissenting).

many courts that adhere to the *M'Naghten* standard of insanity do so because “[t]he criminal law has long been based upon the concept of freedom of choice and adherence to the *M'Naghten* test . . . recognizes that those who are incapable of understanding the wrongfulness of their conduct have no opportunity of choice . . . .”<sup>92</sup>

The “irresistible impulse” test focuses on impairment of defendant’s volitional capabilities rather than the undermining of his cognitive faculties like *M'Naghten*, but the rationale underlying the judicial recognition of the irresistible impulse rule is also tied to the concept of free will. Thus, it is widely believed that the irresistible impulse rule may be invoked only when a defendant demonstrates that the commission of the criminal act “[could not] be resisted or overcome because insanity or mental disease has destroyed the freedom of will, the power of self-control, and the choice of [the defendant’s] actions.”<sup>93</sup> Given that the Model Penal Code test combines the *M'Naghten* rule with the irresistible impulse standard for legal insanity, it is sensible to conclude that the Model Penal Code’s formulation of the defense is connected to free will in much the same manner as the other two tests.

#### E. Duress

A defendant may successfully invoke the defense of duress when he is coerced to engage in wrongful conduct by the use of unlawful force upon the defendant or a third person.<sup>94</sup> According to the Model Penal Code, the force with which the defendant is coerced must be of such a nature that a “person of reasonable firmness . . . would have been unable to resist [it].”<sup>95</sup> A defendant who commits a crime under duress cannot be fairly blamed for engaging in the unlawful act.<sup>96</sup> Therefore, the duress defense functions as an excuse that negates the actor’s culpability without excluding the wrongfulness of his act.<sup>97</sup> The duress excuse is sometimes called the defense of “compulsion,”<sup>98</sup> given that the defendant who acts under duress is “compelled” by threats to engage in an act that he otherwise would not perform. The coercive force that triggers the duress defense must imminently jeopardize the life or limb of the defendant or his family.<sup>99</sup> Therefore, it is generally held that threats to property cannot ground a successful duress defense. The idea underlying this distinction is that society can legitimately expect citizens

---

<sup>92</sup> *People v. Horn*, 205 Cal. Rptr. 119, 129 (Cal. Ct. App. 1984).

<sup>93</sup> *Snider v. Smyth*, 187 F. Supp. 299, 302 (E.D. Va. 1960).

<sup>94</sup> *See, e.g., Smith v. State*, 229 P.3d 221, 226 (Alaska Ct. App. 2010).

<sup>95</sup> MODEL PENAL CODE § 2.09 (2001).

<sup>96</sup> *See FLETCHER, supra* note 22, at 148–49.

<sup>97</sup> Excuse defenses negate the actor’s blameworthiness without negating the wrongfulness of the act. In contrast, justification defenses negate the wrongfulness of the act. Luis E. Chiesa, *Duress, Demanding Heroism, and Proportionality*, 41 VAND. J. TRANSNAT’L L. 741, 746–48 (2008) (discussing the justification/excuse distinction).

<sup>98</sup> *See, e.g., State v. Baker*, 197 P.3d 421, 426–27 (Kan. 2008).

<sup>99</sup> *Id.* at 427.



to sacrifice their property interests in order to avoid wrongfully harming others, whereas it cannot legitimately require its citizens to sacrifice their lives, the lives of their loved ones, or physical integrity in order to prevent unlawful harm to others.<sup>100</sup>

Amongst the many theories that have been advanced to explain the exculpatory force of duress, the two most commonly invoked are the “involuntariness” and the “hard choice” theories of duress. According to the involuntariness theory, “duress exculpates actors whose choice-making capabilities are substantially reduced by the coercive situation that generates the defense.”<sup>101</sup> In these cases, the reduction in the actor’s freedom to choose is so significant that the coerced actor’s choice is in reality “no choice at all”—the duress exculpates the actor because her capacity to choose to do otherwise is “absent” in light of the coercion.<sup>102</sup> According to the hard choice theory of duress, an actor is excused when the coercive situation puts him in the undesirable position of having to decide to preserve either his life or bodily integrity or the fundamental interests of others.<sup>103</sup> Those who face this “do it or else” situation do not act culpably if they choose to harm others because society believes that a reasonable person facing the same hard choice would have acted in the same manner.

Not surprisingly, what the involuntariness and hard choice theories of duress have in common is that the exculpatory force of the defense is in some way tied to the breadth and scope of the actor’s choices. The less choice we have and the harder the choice becomes, the less responsible we are for our “choice.” Thus, the conventional understanding of the duress defense presupposes the existence of an individual endowed with freedom of choice whose will is overborne by coercive threats.<sup>104</sup>

---

<sup>100</sup> See *Prosecutor v. Erdemović*, Case No. IT-96-22-A, Judgment, ¶ 12(c) (Int’l Crim. Trib. for the Former Yugoslavia Oct. 7, 1997) (suggesting that duress creates “no moral alternative” for the person acting under it).

<sup>101</sup> *Chiesa*, *supra* note 97, at 758; see also *People v. Graham*, 129 Cal. Rptr. 31, 32 (Cal. Ct. App. 1976) (establishing that the defendant has the “burden of showing duress” by a raising a reasonable doubt); *State v. Rouleau*, 528 A.2d 343, 350 (Conn. 1987) (“[D]uress may be seen as removing the very basis of criminal culpability.”).

<sup>102</sup> See MARKUS D. DUBBER, *CRIMINAL LAW: MODEL PENAL CODE* 251 (2002).

<sup>103</sup> For an examination and defense of the “hard-choice” theory of excuse and duress, see Stephen J. Morse, *Deprivation and Desert*, in *FROM SOCIAL JUSTICE TO CRIMINAL JUSTICE: POVERTY AND THE ADMINISTRATION OF CRIMINAL LAW* 114, 124–29 (William C. Heffernan & John Kleinig eds., 2000); Michael S. Moore, *Responsibility and the Unconscious*, 53 S. CAL. L. REV. 1563, 1567, 1663–64 (1990).

<sup>104</sup> See, e.g., *Commonwealth v. Wojciechowski*, No. 92456, 2000 WL 537244, at \*3 (Mass. Super. Ct. Feb. 25, 2000) (stating that the prosecution must “prove[] beyond a reasonable doubt that the crime was the result of that person’s free will and not the result of duress or coercion such as would destroy that person’s free will”); see also *Minton v. State*, 305 S.E.2d 812, 814 (Ga. Ct. App. 1983) (stating that defendant’s conduct is only criminal if he decided to commit the crime “as a result of his own free will . . . not the result of duress or coercion which would overcome his free will”).

*F. Entrapment*

An actor may sometimes invoke police entrapment as a defense to criminal liability. There are two versions of the entrapment defense. According to the “objective” version of the defense, a defendant who the police instigated to commit a crime may successfully plead entrapment “if the law enforcement conduct is likely to induce a *normally law-abiding person* to commit the offense.”<sup>105</sup> Therefore, the objective approach to entrapment “focus[es] on the conduct of the police and [is] not concerned with the defendant’s prior criminal activity or other indicia of a predisposition to commit crime.”<sup>106</sup> In contrast, the subjective version of entrapment focuses on the defendant’s readiness to commit the crime rather than on the outrageousness of the governmental conduct.<sup>107</sup> More specifically, the subjective version of the defense can only be invoked if the police induced a person to commit an offense that he was not predisposed to commit prior to being approached by governmental authorities.<sup>108</sup> Therefore, “[t]he subjective defense of entrapment succeeds only if the government, not the accused, is the source of the criminal design.”<sup>109</sup> The defense, however, “fails if the accused is previously disposed to commit the crime, and the government merely facilitates or assists the criminal scheme.”<sup>110</sup> Although the objective version of the defense has garnered considerable support amongst commentators, and was adopted by the drafters of the Model Penal Code, the subjective approach to entrapment continues to be applied in a majority of American jurisdictions.<sup>111</sup>

With the understanding that the subjective version of entrapment allows the actor to invoke the defense only if he was not predisposed to commit the crime before being approached by the police, the defense fails if his decision to commit the criminal act was a product of his own volition. On the other hand, the defendant can successfully invoke entrapment if he proves that the police caused him to decide to commit the offense. Whether the entrapment defense bars criminal liability thus hinges on whether the defendant can be said to have freely willed the commission of the crime. More specifically, the entrapment defense will succeed if the defendant did not freely will the commission of the crime because the police caused him to engage in the wrongful act. However, the defendant’s entrapment claim will fail if he voluntarily decided to commit the crime.

It has been held that “[o]nce the entrapment defense is raised, the State bears the burden of showing that the defendant was predisposed to commit the crime and that the level of police activity did not persuasively affect the free will of the

---

<sup>105</sup> *People v. Watson*, 990 P.2d. 1031, 1032 (Cal. 2000).

<sup>106</sup> *Commonwealth v. Weiskerger*, 554 A.2d 10, 13 (Pa. 1989).

<sup>107</sup> *See, e.g., Hernández v. State*, 17 So. 3d 748, 750–51 (Fla. Dist. Ct. App. 2009).

<sup>108</sup> *See, e.g., Jacobson v. United States*, 503 U.S. 540, 548–50 (1992).

<sup>109</sup> *One Way Fare v. Dep’t of Consumer Prot.*, 901 A.2d 1246, 1249 (Conn. App. Ct. 2006).

<sup>110</sup> *Id.*

<sup>111</sup> 2 WAYNE R. LAFAVE, *SUBSTANTIVE CRIMINAL LAW* § 9.8(b) (2d ed. 2003).

accused.”<sup>112</sup> Therefore, the determinative inquiry in entrapment cases is whether “the agents overpowered the free will of the defendant and caused him to commit” an offense that he was not predisposed to commit.<sup>113</sup> Therefore, the availability of the subjective version of the entrapment defense, as with many other criminal defenses, depends on considerations of free will and voluntariness.

### III. DO WE HAVE FREE WILL? SCIENTIFIC AND PHILOSOPHICAL PERSPECTIVES

Free will is as essential to explaining our current practices of blaming and punishing as it is to understanding *Candide*’s view of human nature.<sup>114</sup> But do we really have the sort of free will that allows us to make sense of *Candide*’s conception of evil and in turn furnishes the foundations for our system of criminal justice? At first glance, this appears to be a silly question. After all, humans share a deeply held belief that what we do is usually a product of what we consciously want to do. However, as history demonstrates, widely shared beliefs often turn out to be false. For thousands of years, humans believed that the world was flat and that earth stood still while the sun and the planets revolved around it. Could it be that, despite appearances to the contrary, the widely shared belief that humans are endowed with free will is as false as these other once widely held beliefs are? As it turns out—and as the rest of this section explains—it is still very much an open question whether we do in fact have the kind of free will that undergirds our criminal laws.

#### A. Causal Determinism and the Free Will Problem

One of the problems frequently presented in debates about free will is that the meaning of “free will” often remains undefined or is only vaguely or confusingly fleshed out. In order to avoid such confusion, fruitful discussions about the nature and importance of free will must begin by defining the term. The problem is that there are many plausible ways to define free will.<sup>115</sup> So how can one defend choosing one definition of free will over another? A nonarbitrary and promising way to choose amongst competing definitions of free will is to select the definition that sheds more light on the question addressed by the author. It thus makes sense to define free will for the purposes of this Article in a way that illuminates the

---

<sup>112</sup> *Jordan v. State*, 692 N.E.2d 481, 484 (Ind. Ct. App. 1998).

<sup>113</sup> *United States v. Jaca-Nazario*, 521 F.3d 50, 58 (1st Cir. 2008) (quoting *United States v. Barbour*, 393 F.3d 82, 86 (1st Cir. 2004)).

<sup>114</sup> See *supra* notes 1–21 and accompanying text.

<sup>115</sup> Thomas Hobbes, for example, defined free will as acting in an uncoerced fashion or, more specifically, acting in the absence of external impediments. THOMAS HOBBS, *LEVIATHAN* 86 (1904). In contrast, Harry Frankfurt defines free will as an action performed under circumstances in which the actor identifies with the volition that motivated his act. Harry G. Frankfurt, *Freedom of the Will and the Concept of a Person*, in *FREE WILL* 81–95 (Gary Watson ed., 1982). Other philosophers define free will differently. See generally Taylor & Dennett, *supra* note 15, at 258 (stating multiple definitions of free will).

connection between free will and criminal liability. Free will is relevant to criminal liability because blame is typically a prerequisite for the imposition of punishment. It is generally believed that an actor can be blamed for committing an offense only if he freely willed to engage in conduct constitutive of the offense. As a result, it makes sense to define free will for the purposes of this Article as *the degree of freedom that makes judgments of blame and attributions of moral responsibility possible*.

Once free will is defined in this manner, it is easy to see why many contemporary scientists and philosophers believe that free will is threatened by causal determinism.<sup>116</sup> Causal determinism is the belief that everything that happens in the universe, including human conduct, is the product of everything that has happened in the past in combination with the operation of natural laws.<sup>117</sup> Contrary to what some have argued, causal determinism does not lead to the conclusion that human desires and intentions cannot impact future events because “what will be, will be” regardless of what actors wish or desire to do. Aristotle, for example, claimed that if causal determinism obtains there would be no need to deliberate or think about our future actions because whatever will happen will in fact happen regardless of what we think or do.<sup>118</sup> Determinists do not necessarily share this thesis, because they believe that human conduct is causally determined by myriad factors, including intentions, motives, and desires. Whether a Mayweather-Pacquiao boxing match will take place in the future is causally conditioned by a series of factors, including, but not limited to, the desires of Mayweather and Pacquiao. Determinists, of course, also believe that such intents, motives, and desires are causally determined by other factors, such as the actor’s upbringing and her genetic makeup.<sup>119</sup> This, however, does not commit the determinist to believing that human beliefs and desires cannot impact future states of affairs, for the determinist thesis affirms the causal role played by such subjective mental states.

There are at least four reasons that suggest the thesis of causal determinism ought to be taken seriously. First, there is scientific support for the view that the behavior of macroscopic objects (including human beings) is causally determined by the confluence of the past and natural laws. The conventional view is that classical physics is mostly a deterministic affair.<sup>120</sup> There are, however, instances

---

<sup>116</sup> See generally THOMAS PINK, *FREE WILL: A VERY SHORT INTRODUCTION* 14–16 (2004) (explaining the concept of causal determinism and demonstrating that the existence of human free will is likely dependent on the truth of causal determinism).

<sup>117</sup> John Martin Fischer, *Compatibilism*, in *FOUR VIEWS ON FREE WILL* 44, 54.

<sup>118</sup> ARISTOTLE, *DE INTERPRETATIONE*, ch. 9, §§ 18<sup>b</sup>31–3 (John Lloyd Ackrill trans. 1975).

<sup>119</sup> See PINK, *supra* note 116, at 7; John Martin Fischer, Robert Kane, Derek Pereboom & Manuel Vargas, *A Brief Introduction to Some Terms and Concepts*, in *FOUR VIEWS ON FREE WILL* 1, 3.

<sup>120</sup> See generally JOHN EARMAN, *A PRIMER ON DETERMINISM* (Robert S. Butts ed., 1986).

in which the deterministic nature of classical physics breaks down.<sup>121</sup> Contrarily, quantum mechanics is conventionally viewed as mostly governed by indeterminism. Nevertheless, there might be instances in which determinism seeps into quantum mechanics.<sup>122</sup> In any case, there seems to be significant scientific support for the view that some have called “adequate determinism,” which holds that indeterministic quantum effects are mostly negligible at the macroscopic level. Second, neuroscientific studies, such as the ones famously carried out by Benjamin Libet, suggest that human conduct is determined by unconscious processes that are not within the control of the actor.<sup>123</sup> Third, biological studies demonstrate that some aspects of human behavior are largely determined by our genetic makeup.<sup>124</sup> Finally, several psychological studies hint at the possibility that many aspects of human behavior are determined to a significant extent by environmental factors.<sup>125</sup> None of these reasons alone are enough to establish the truth of causal determinism. Nevertheless, the combination of all of these factors does, at the very least, cast serious doubt on whether human conduct can be traced back to indeterministic processes. In fact, the strongest evidence to date in favor of the conclusion that human conduct is *not* causally determined is the robust, but scientifically unsupported, intuition that we control our destinies in a way that belies the thesis of determinism.

Most people believe that actors can be blamed or praised for what they do only if they have the ability to choose to act differently. Philosophers call this the principle of alternate possibilities.<sup>126</sup> Causal determinism threatens the principle of alternate possibilities because it suggests that, given the fixity of the past and the immutability of natural laws, humans lack control over the factors that shape their conduct. As a result, some philosophers—called incompatibilists—have concluded that free will is incompatible with causal determinism.<sup>127</sup> “Libertarian” incompatibilists believe that we possess free will because the thesis of causal determinism is likely false.<sup>128</sup> In contrast, “hard incompatibilists” believe that if

---

<sup>121</sup> *Id.*

<sup>122</sup> See Robert C. Bishop, *Determinism and Indeterminism*, in 3 *ENCYCLOPEDIA OF PHILOSOPHY* 29, 32–33 (Donald M. Borchert ed., 2d ed. 2005).

<sup>123</sup> Benjamin Libet, *Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action*, in 8 *BEHAV. & BRAIN SCI.* 529, 529 (1985) (describing the results of a study of electrophysical “readiness potentials” that precede voluntary acts).

<sup>124</sup> For an account of how genes influence behavior, see generally MICHAEL C. RUTTER, *GENES AND BEHAVIOR: NATURE-NURTURE INTERPLAY EXPLAINED* (2006) (asserting the importance of genetic science and outlining scientific theories regarding the influence of genes on behavior).

<sup>125</sup> See, e.g., Nadia Gamefski & Sjoukje Okma, *Addiction-Risk and Aggressive/Criminal Behaviour in Adolescence: Influence of Family, School and Peers*, 19 *J. ADOLESCENCE* 503–12 (1996) (detailing one study of the effects of family, school, peers, and other environmental factors on secondary school students in the Netherlands).

<sup>126</sup> The principle of alternate possibilities was first explored as such by Harry Frankfurt in *Alternate Possibilities and Moral Responsibility*, 66 *J. PHIL.* 829 (1969).

<sup>127</sup> Fischer, Kane, Pereboom & Vargas, *supra* note 119, at 1, 3.

<sup>128</sup> *Id.*

causal determinism is true, we should abandon the idea of free will. They also argue that free will is incompatible with indeterminism.<sup>129</sup> The so-called “compatibilists” believe that we have the sort of free will that underlies our judgments of blame and praise even if causal determinism obtains.<sup>130</sup> The remainder of Part III explores in more detail the views held by libertarians, hard incompatibilists and compatibilists.

### B. *Libertarianism*

Libertarians argue that belief in free will is not compatible with a deterministic account of human behavior. According to libertarianism, true freedom of the will requires actual and accessible alternative possibilities.<sup>131</sup> That is, conduct is free in the sense required for moral responsibility if, and only if, the actor could have chosen to do otherwise.<sup>132</sup> An actor has access to alternative possibilities if there are different possible courses of conduct available to him and it is ultimately up to him to decide which one of the options to pursue. Libertarians concede that actors would not have access to these alternative possibilities if the thesis of causal determinism were true.<sup>133</sup> Nevertheless, they argue that humans do have access to the alternative possibilities that free will requires because causal determinism is false.<sup>134</sup> Thus, libertarians maintain that irreducibly indeterminate processes ultimately produce human conduct.<sup>135</sup> As a result, the soundness of the libertarian thesis depends on whether a plausible scientific case can be made that undetermined forces cause human conduct, unlike the behavior of most other things in the universe.

The distinguished libertarian philosopher Robert Kane has come up with an ingenious way of defending the thesis that human behavior can be traced back to indeterministic processes. Kane mounts his defense of indeterminism by borrowing heavily from the field of quantum mechanics.<sup>136</sup> Quantum mechanics are the laws of physics that apply to the behavior of atomic and subatomic particles.<sup>137</sup> What makes quantum mechanics interesting—and what made Einstein cringe whenever he talked about the implications of the theory—is that atomic and subatomic particles behave quite differently than larger particles.<sup>138</sup> Whereas the

---

<sup>129</sup> *Id.*

<sup>130</sup> *Id.*

<sup>131</sup> Laura Waddell Ekstrom, *Libertarianism and Frankfurt-Style Cases*, in THE OXFORD HANDBOOK OF FREE WILL 309, 321.

<sup>132</sup> *Id.*

<sup>133</sup> *Id.*

<sup>134</sup> Robert Kane, *Libertarianism*, in FOUR VIEWS ON FREE WILL 5, 7.

<sup>135</sup> *Id.* at 26–28.

<sup>136</sup> *Id.* at 29.

<sup>137</sup> Jenann Ismael, *Quantum Mechanics*, THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Sep. 1, 2009), <http://plato.stanford.edu/entries/qm>.

<sup>138</sup> Einstein was troubled by the nondeterminist nature of quantum mechanics. This is what led him to famously reject the implications of quantum mechanics by succinctly

behavior of macroscopic particles appears to be fully determined by causal laws, quantum mechanics suggests that the behavior of microscopic particles is not causally determined in the same way.<sup>139</sup> More specifically, quantum mechanics holds that we can attempt to predict the behavior of subatomic particles, but we can never be certain of how these particles are actually going to behave.<sup>140</sup> Kane puts this insight to good use by positing that the neuron firings that ultimately determine our conduct are the product of processes that originate at the subatomic level and, therefore, are not fully determined by the past and the operation of natural laws.<sup>141</sup> This allows Kane to suggest that, in light of quantum mechanics and the possibility that such processes undergird the workings of the human brain, causal determinism may very well prove to be false, at least insofar as it purports to explain human behavior.<sup>142</sup>

### C. *Hard Incompatibilism*

Hard incompatibilists, like libertarians, believe that free will is incompatible with causal determinism.<sup>143</sup> However, unlike libertarians, hard incompatibilists argue that free will is also incompatible with indeterminism. Therefore, hard incompatibilists suggest that belief in free will should be abandoned regardless of whether human conduct is causally determined.<sup>144</sup> There are at least two reasons that lead hard incompatibilists to reject the existence of the sort of free will that undergirds our judgments of blame and praise. First, if causal determinism is true, it would seem that humans lack the ability to do something that seems to be essential to their moral responsibility—the capacity to do otherwise or, at the very least, the capacity for the agent to be the ultimate source or originator of her decision to act in a particular way.<sup>145</sup>

Furthermore, proponents of hard incompatibilism argue that free will is undermined even if it turns out that human conduct is the product of indeterministic processes. They do so by contending that it only makes sense to blame or praise us for our acts if we are the ultimate source of the desires, motives and intentions that shape our conduct.<sup>146</sup> If, however, it turns out that our beliefs, intentions, and motives are the product of indeterministic processes that are not

---

stating that God “does not play dice” with the universe. MAX BORN, *THE BORN-EINSTEIN LETTERS: FRIENDSHIP, POLITICS AND PHYSICS IN UNCERTAIN TIMES* 152 (Irene Born trans., 2d ed. 2005).

<sup>139</sup> Ismael, *supra* note 137.

<sup>140</sup> See Diana Buchwald & Kip Thorne, *Preface* to MAX BORN, *THE BORN-EINSTEIN LETTERS: FRIENDSHIP, POLITICS AND PHYSICS IN UNCERTAIN TIMES* vii, xxii; Ismael, *supra* note 137.

<sup>141</sup> Kane, *supra* note 134, at 7.

<sup>142</sup> *Id.* at 29–30.

<sup>143</sup> Fischer, Kane, Pereboom & Vargas, *supra* note 119, at 3.

<sup>144</sup> Derek Pereboom, *Hard Incompatibilism*, in *FOUR VIEWS ON FREE WILL*, 85.

<sup>145</sup> *Id.* at 97.

<sup>146</sup> *Id.* at 97–98.

governed by causal laws, it would seem that our decisions and acts would be the product of random and haphazard events that we cannot control and thus cannot ground the sort of freedom that undergirds our judgments of desert.<sup>147</sup>

A famous recent argument in favor of the hard incompatibilist claim that determinism is incompatible with free will is Derk Pereboom's "four case manipulation argument."<sup>148</sup> The first case put forth in the argument presents us with an evil neuroscientist who creates a human that can be manipulated directly by radio controls and who is in fact manipulated by the scientist in a way that causes him to murder Ms. White.<sup>149</sup> In the next case, the neuroscientist creates a human who cannot be directly manipulated like in the first case, but who has a brain that is programmed in a way that causally determines him to murder Ms. White.<sup>150</sup> The third case tells the story of an ordinary human who has been rigorously trained since he was an infant by his parents and his community in a way that causes him to have certain inclinations and character traits that, in turn, cause him to murder Ms. White.<sup>151</sup> Finally, in the fourth case we have a normal human being living in a universe where causal determinism obtains who is caused to murder Ms. White as a result of the interplay of events that happened in the past (e.g., his upbringing) and the operation of natural laws.<sup>152</sup> Pereboom then contends that our intuitions strongly suggest that the agent is not responsible in cases one and two.<sup>153</sup> If so, he argues that, given the similarities between cases two and three, and cases three and four, we should also conclude that the agents in cases three and four are not responsible, because the ultimate causes of the agent's action in all of these instances can be traced back to events beyond the agent's control.<sup>154</sup> If one accepts this conclusion, then causal determinism is incompatible with free will, because the agent in case four is like any human being who acts in a causally determined universe.<sup>155</sup>

#### D. Compatibilism

The so-called "compatibilists" have staked out an increasingly important stance on the free will debate. The adherents of this theory do not deny the plausibility of the thesis of causal determinism. As a matter of fact, some compatibilists believe that causal determinism is likely true.<sup>156</sup> Nevertheless, the

---

<sup>147</sup> *Id.* at 109.

<sup>148</sup> Pereboom, *supra* note 30, at 110–17.

<sup>149</sup> *Id.* at 112–13.

<sup>150</sup> *Id.* at 113–14.

<sup>151</sup> *Id.* at 114–15.

<sup>152</sup> *Id.* at 115–16.

<sup>153</sup> *Id.* at 112–13.

<sup>154</sup> *Id.* at 115–17.

<sup>155</sup> *Id.* at 116.

<sup>156</sup> See, e.g., DANIEL C. DENNETT, FREEDOM EVOLVES 305 (2003) (claiming that "all the products of our brains have . . . been designed . . . by physical processes in which no exemption from causality can be discerned").



compatibilist argument suggests that humans possess the sort of free will that underlies judgments of moral responsibility regardless of whether it turns out that causal determinism is true.<sup>157</sup> As a result, the defenders of compatibilism believe that free will is not threatened at all by causal determinism.<sup>158</sup> Although compatibilists agree that the freedom of will that is needed to justify our practices of blaming and punishing is compatible with the truth of causal determinism, there is significant disagreement regarding the reasons that justify this conclusion. While many compatibilists espouse so-called “mesh” accounts of compatibilism,<sup>159</sup> others endorse what can be dubbed a “reactive attitudes” account of compatibilism.<sup>160</sup> Furthermore, some theorists refuse to label themselves as “compatibilists” although their accounts of free will and its relationship with moral responsibility seem to share important features of the compatibilist stance.<sup>161</sup>

According to mesh accounts of compatibilism, an actor’s will is free if there is an appropriate “mesh or connection” between her choices and her desires and preferences.<sup>162</sup> Perhaps the most influential mesh account of compatibilism is Harry Frankfurt’s.<sup>163</sup> The starting point of Frankfurt’s theory is that humans have both first and second order volitions.<sup>164</sup> First order volitions are volitions to bring about an action or a certain state of affairs,<sup>165</sup> such as my desire to eat the Cheesecake Factory’s sinful red velvet cheesecake. In contrast, second order volitions are volitions about first order volitions,<sup>166</sup> such as my desire to not desire to eat the aforementioned red velvet cheesecake. According to Frankfurt, an actor acts freely only if his second order volition is aligned with his first order volition,<sup>167</sup> such as when I eat the red velvet cheesecake in circumstances in which I desire to eat the red velvet cheesecake and I desire to act in accordance with my desire to eat the red velvet cheesecake. Contrarily, an actor does not act freely when her second order volition is not consistent with her first order volition, such as when I eat a red velvet cheesecake that I desired to eat although I did not desire to act in accordance with my desire to eat the dessert (because I am on a diet, for example).<sup>168</sup> Frankfurt’s theory is compatibilist because an agent can act freely under her account as long as the actor’s second order volition is connected in a certain way with her first order volition (i.e., she eats the cheesecake and she desires to act in accordance with her desire to eat the cheesecake) even if she

---

<sup>157</sup> Fischer, Kane, Pereboom & Vargas, *supra* note 119, at 3.

<sup>158</sup> PINK, *supra* note 116, at 18–19.

<sup>159</sup> Haji, *supra* note 31, at 210–25.

<sup>160</sup> *Id.* at 203–04.

<sup>161</sup> *See, e.g.*, Fischer, *supra* note 117, at 71–78.

<sup>162</sup> Haji, *supra* note 31, at 210.

<sup>163</sup> FRANKFURT, *supra* note 115.

<sup>164</sup> *Id.* at 7.

<sup>165</sup> *Id.*

<sup>166</sup> *Id.*

<sup>167</sup> *Id.* at 15.

<sup>168</sup> *Id.*

cannot do anything other than what she does in light of the causally determined nature of the universe (i.e., she cannot decide to abstain from eating the cake).<sup>169</sup>

Several decades ago, P. F. Strawson advanced another influential account of compatibilism. Strawson's theory begins by positing the importance of "reactive attitudes" to the human experience.<sup>170</sup> Reactive attitudes are reactions that are essential to establishing and maintaining personal relationships.<sup>171</sup> These reactions include, but are not limited to, feelings of resentment, gratitude, forgiveness, anger, and love.<sup>172</sup> According to Strawson, we act freely and, therefore, are morally responsible for our conduct when we behave in a way that generates reactive attitudes in others.<sup>173</sup> Strawson believes that this type of freedom and moral responsibility is not undermined by the truth of causal determinism.<sup>174</sup> More specifically, he argues that we must continue to assume that humans have free will even if determinism obtains.<sup>175</sup> Regardless of how scientifically plausible the incompatibilist argument may seem, we should not assume that we lack free will, for doing so would jeopardize the reactive attitudes and, thus, the kind of interpersonal relationships that are essential to our lives.<sup>176</sup>

#### *E. Semi-Compatibilism*

John Martin Fischer and Mark Ravizza advanced a highly influential solution to the free will problem that they dubbed "semicompatibilism."<sup>177</sup> Semicompatibilism accepts as plausible the view that free will requires access to alternative possibilities and that causal determinism is incompatible with free will because it negates access to such alternative possibilities.<sup>178</sup> Nevertheless, semicompatibilists believe that moral responsibility does not require access to alternative possibilities and is thus compatible with causal determinism.<sup>179</sup> The gist of Fischer and Ravizza's theory is that a person acts in a morally responsible way if her conduct is responsive to reason.<sup>180</sup> Under this account, compulsives and the mentally ill often do not act in a morally responsible manner because they are not

---

<sup>169</sup> *Id.* at 19. Therefore, Frankfurt (famously) rejects the principle of alternate possibilities as a prerequisite for free will and moral responsibility. *See* Frankfurt, *supra* note 126.

<sup>170</sup> Strawson, *supra* note 18, at 47–50.

<sup>171</sup> *Id.*

<sup>172</sup> *Id.* at 48.

<sup>173</sup> *Id.* at 50–55.

<sup>174</sup> *See id.* at 55.

<sup>175</sup> *Id.* at 55–56.

<sup>176</sup> *Id.*

<sup>177</sup> *See generally* JOHN MARTIN FISCHER & MARK RAVIZZA, RESPONSIBILITY AND CONTROL 10 (1999) (presenting a general theory that designates the circumstances where individuals are morally responsible for their actions and omissions).

<sup>178</sup> *See* Fischer, *supra* note 117, at 71–74.

<sup>179</sup> *Id.* at 78, 80.

<sup>180</sup> *See id.* at 78–79.

responsive to the sort of rational considerations that guide the conduct of the majority of humans. Furthermore, Fischer and Ravizza contend that an individual is morally responsible for her conduct only if her action is triggered by a decision or volition that can be properly characterized as belonging to the agent.<sup>181</sup> Thus, ascriptions of moral responsibility are warranted in cases where the behavior in question was caused by the actor's *own* reason-responsive process of deliberation, as opposed to it being caused by a non-reason-responsive decisional process (the mentally ill) or a reason-responsive deliberative process that does not belong to the agent (thought processes induced by hypnosis).<sup>182</sup> Finally, Fischer and Ravizza argue that these attributions of moral responsibility are justified even if causal determinism obtains because determinism is compatible with the view that human conduct can be caused by the actor's reason-responsive thought processes.

#### IV. FREE WILL IN CRIMINAL THEORY

As Part II of this Article demonstrates, courts and commentators traditionally view free will as essential to criminal responsibility. Nevertheless, many, if not most, criminal theorists have little to say about whether our practices of blaming and punishing are undermined by the thesis of causal determinism. Of the few criminal law scholars who discuss the free will problem, most espouse views that mirror the theories of free will and moral responsibility discussed in Part III. Thus, some criminal theorists, like George Fletcher, defend a libertarian account of free will.<sup>183</sup> Others, like Michael Moore, Stephen Morse, and most continental criminal scholars, defend a compatibilist account of free will.<sup>184</sup> Larry Alexander, for his part, mysteriously declares himself to be agnostic with regard to the existence of free will and its compatibility with causal determinism. Curiously, though some scholars such as Anders Kaye have flirted with hard incompatibilist accounts of free will,<sup>185</sup> no criminal theorist has comprehensively defended the view that causal determinism is likely true and that it is incompatible with free will and moral responsibility. This Part seeks to explain, in more detail, the way these and other criminal theorists have approached the free will problem.

##### A. *George Fletcher's Chomskyan-Inspired Libertarianism*

Distinguished criminal law scholar George Fletcher bravely tackled the free will problem in his recently published *Grammar of Criminal Law*.<sup>186</sup> The problem is discussed in the context of analyzing the nature and scope of the criminal law's

---

<sup>181</sup> See *id.* at 79.

<sup>182</sup> *Id.* at 79–80 (stating that “one relevant feature to the actual-sequence mechanism is that it must be in some appropriate way responsive to reasons”).

<sup>183</sup> See *supra* notes 131–142 and accompanying text.

<sup>184</sup> See *supra* notes 156–176 and accompanying text.

<sup>185</sup> See *supra* notes 143–155 and accompanying text.

<sup>186</sup> See FLETCHER, *supra* note 33, at 273–81.

act requirement.<sup>187</sup> According to Fletcher, the relevance of free will to criminal law is that conduct satisfies the act requirement if, and only if, it is the product of the effort or determination of the actor.<sup>188</sup> As Fletcher correctly points out, it is difficult to defend this way of construing the act requirement if the thesis of causal determinism turns out to be true.<sup>189</sup> Fletcher appears to be quite troubled by this problem, given that conduct is not really the product of the actor's effort or determination if it is fully caused by the interplay of past events and the operation of natural laws. Therefore, Fletcher appears to believe that criminal responsibility is incompatible with causal determinism. He thus attempts to solve the free will problem by arguing that causal determinism is false, at least insofar as it purports to explain the inner workings of human behavior.<sup>190</sup>

Fletcher advances two arguments that he believes undermine the thesis of causal determinism. First, relying on Noam Chomsky's theory of language, Fletcher argues that the number of sentences that we command in natural language is infinite.<sup>191</sup> This leads him to conclude that human beings could not be predetermined or programmed to invent and understand new phrases because the number of phrases that could be invented are infinite and thus cannot be determined or programmed *ex ante*.<sup>192</sup> Fletcher's second argument is based on John Searle's famous Chinese Room argument.<sup>193</sup> The Chinese Room argument asks us to imagine a person inside a room with a set of instructions in English that allow him to coherently answer questions posed to him in Chinese although he does not read or understand Chinese.<sup>194</sup> Properly understood, the argument is supposed to show that computers that follow algorithms, much like the person inside the Chinese room, cannot possibly understand the meaning of the answers it provides to the problems with which it is posed.<sup>195</sup> Fletcher believes that the Chinese room argument casts doubt on the plausibility of the thesis of causal determinism because it shows that it cannot be determined in advance that human beings understand the meaning of language, because Searle demonstrates that no computer could be programmed in advanced to achieve this feat.<sup>196</sup> This leads Fletcher to conclude that actors are endowed with free will because human conduct is unprogrammed, infinitely variable, and thus not causally determined.<sup>197</sup>

---

<sup>187</sup> *Id.* at 266–97.

<sup>188</sup> *Id.* at 273.

<sup>189</sup> *Id.* at 275.

<sup>190</sup> *Id.* at 278 (although Fletcher never expressly states that he is an “incompatibilist” he defends what he calls an “antideterminist” position, which, as the term implies, holds that determinism is false).

<sup>191</sup> *Id.* at 278.

<sup>192</sup> *Id.* at 279.

<sup>193</sup> *Id.* at 279. For Searle's own discussion of the Chinese Room argument, see John R. Searle, *Minds Brains and Programs*, 3 BEHAV. & BRAIN SCI. 417 (1980).

<sup>194</sup> FLETCHER, *supra* note 33, at 279.

<sup>195</sup> *Id.*

<sup>196</sup> *Id.*

<sup>197</sup> *Id.*

## B. *Compatibilism in Criminal Theory*

### 1. *Stephen Morse and Michael Moore's Compatibilism*

Stephen Morse is one of only a handful of criminal theorists that have seriously tackled the free will problem. His views have proved to be quite influential amongst criminal scholars. Morse, unlike Fletcher, believes that causal determinism does not threaten free will and criminal responsibility.<sup>198</sup> Morse also disagrees with Fletcher regarding the plausibility of the thesis of causal determinism. While Fletcher believes that causal determinism is likely to be false, Morse believes that causal determinism is probably true.<sup>199</sup> Nevertheless, Morse argues that the sort of free will that is essential to moral responsibility is compatible with determinism.<sup>200</sup> Thus, he defends a compatibilist account of free will. Furthermore, Morse argues that most of the criminal law doctrines that are currently in place can be recast in compatibilist terms without effecting major changes in criminal law jurisprudence.<sup>201</sup>

Morse's compatibilist views closely resemble those espoused by philosopher Daniel Dennett. Dennett argues that the sort of freedom that undergirds our practices of blaming consists in the ability to give reasons in favor and against engaging in a certain conduct.<sup>202</sup> Similarly, Morse suggests that the freedom that is required by moral responsibility "is dependent primarily on the agent's general capacity to grasp and be guided by reason."<sup>203</sup> Therefore, Morse believes that an actor is morally responsible for his choices as long as he has the capacity to be guided by reason. Contrarily, he considers that an actor is not responsible if her capacity to be guided by reasons is diminished or nonexistent.<sup>204</sup> According to Morse, the advantage of his account is that it allows us to continue holding people morally responsible for their conduct even if it turns out that determinism is true. The fact that everything that takes place in the universe is causally determined does not impinge upon the uniquely human capacity to rationally mull over our choices and plan our conduct accordingly. Thus, Morse suggests that the sort of

---

<sup>198</sup> See, e.g., Stephen J. Morse, *Determinism and the Death of Folk Psychology: Two Challenges to Responsibility from Neuroscience*, 9 MINN. J. L. SCI. & TECH. 1, 14 (2008).

<sup>199</sup> Stephen J. Morse, *Criminal Responsibility and the Disappearing Person*, 28 CARDOZO L. REV. 2545, 2551 (stating that "[c]riminal theorists should unflinchingly accept that we live in a causal universe—that people are part of that natural, causal universe—and consider what follows from these truths").

<sup>200</sup> *Id.* at 2552–53.

<sup>201</sup> See *id.* at 2551–53.

<sup>202</sup> DENNETT, *supra* note 156, at 251 (stating that the freedom that sets us apart from nonhuman animals is the ability to "engage in the practice of asking, and giving, reasons").

<sup>203</sup> Stephen J. Morse, *Reason, Results, and Criminal Responsibility*, 2004 U. ILL. L. REV. 363, 441.

<sup>204</sup> *Id.*

considerations that undergird his conception of free will, are unaffected by the possible truth of determinism.<sup>205</sup>

Like Stephen Morse, distinguished criminal theorist Michael Moore believes that the sort of free will that undergirds our judgments of praise and blame is compatible with determinism.<sup>206</sup> Moore begins his defense of compatibilism by pointing out that actors should be held responsible for their conduct even if the conduct has been caused by factors that the actor does not control—contrary to what some criminal scholars have argued.<sup>207</sup> He thus suggests that responsibility does not hinge on being free from causal processes.<sup>208</sup> Rather, he argues that we ought to excuse certain conduct whenever it is performed under compulsion.<sup>209</sup>

Compulsion differs from causation in that compulsion interferes with one's ability to do what is required by reason and morality, whereas causation produces conduct without interfering with the actor's ability for practical reasoning.<sup>210</sup> Compulsion can be internal, such as a schizophrenic's compulsion to kill his daughter in order to satisfy the will of the voices he hears inside his head. Or it can be external, such as when an actor feels compelled to comply with the unlawful demands of a gun-wielding assailant in order to keep himself free from harm. Moore then suggests that actors are morally responsible for their acts as long as they are not the product of compulsion, even if forces that actors cannot control cause the acts.<sup>211</sup> At first glance, this account of moral responsibility is compatible with the truth of causal determinism, because responsibility would depend on the absence of compulsion rather than on the absence of causation.

Although Moore places much emphasis on the distinction between causation and compulsion, he acknowledges that the distinction is not in and of itself enough to disprove the thesis that moral responsibility is incompatible with determinism. More specifically, Moore admits that there is a strong intuition that an actor is responsible only if he could have acted otherwise and that the truth of causal determinism calls into question whether humans actually have this ability.<sup>212</sup> Moore's response to this objection borrows much from Strawson's "reactive attitudes" account of compatibilism.<sup>213</sup> Moore agrees with Strawson's contention that reactive attitudes such as blame, praise, resentment, and gratitude are an

---

<sup>205</sup> See *id.* at 439.

<sup>206</sup> Moore, *supra* note 11, at 1140 (stating that "people can exercise their will even though their actions are caused").

<sup>207</sup> *Id.*

<sup>208</sup> *Id.*

<sup>209</sup> *Id.*

<sup>210</sup> *Id.* at 1129–30.

<sup>211</sup> *Id.* at 1132.

<sup>212</sup> *Id.* at 1141 (stating that "[e]ven after the conceptual confusions are cleared away, the causal theorist has a moral argument: if a person could not help doing what he did, he is not morally responsible for it. This is true even if his behavior was an action, and an action not covered by the conventional moral excuses").

<sup>213</sup> See *supra* notes 170–176 and accompanying text.

essential part of the human experience.<sup>214</sup> These attitudes are so essential to our lives that they give meaning to most of our valuable interpersonal relationships, including those that we have with our children and friends.<sup>215</sup> Furthermore, Moore believes that reactive attitudes play an important role in guiding our behavior.<sup>216</sup> Forsaking such attitudes would thus lead to falsifying much of our moral life. As a result, Moore contends that we should not renounce to the sort of free will that makes experiencing these attitudes possible, even if it turns out that the thesis of causal determinism is true.<sup>217</sup>

## 2. *Compatibilism in Continental Criminal Theory*

Like their Anglo-American counterparts, many continental criminal law scholars acknowledge that the plausibility of the thesis of causal determinism threatens to undermine the kind of freedom that undergirds conventional theories of criminal responsibility. Claus Roxin—one of Europe’s leading criminal theorists—believes that he has found a way out of the free will problem. Roxin argues that an individual acts freely, and is thus morally and criminally liable, when he is in a position to understand and take into account the applicable legal rules that are intended to govern his conduct.<sup>218</sup> Furthermore, he argues that we ought to assume—without being able to prove—that the individual who is in such a position is also able to conform his conduct to the mandates of the law.<sup>219</sup> This, however, does not mean that Roxin argues for an indeterminist account of freedom of will and responsibility.<sup>220</sup> As a matter of fact, Roxin believes that his view is compatible with determinist accounts of human conduct.<sup>221</sup> More specifically, he suggests that whether humans are capable of acting otherwise is irrelevant to his views on freedom and responsibility because actors who are capable of understanding and taking into account applicable legal rules are treated *as if* they were free to act otherwise, regardless of whether they were actually capable of doing so.<sup>222</sup> Thus, Roxin contends that assuming that humans are endowed with the freedom to choose amongst alternative courses of action is socially valuable regardless of whether epistemological or scientific inquiry casts doubt on whether such a faculty actually exists.<sup>223</sup>

Another compatibilist account of free will, influential amongst continental criminal theorists, is the one advanced by Gunther Jakobs. Jakobs’ compatibilist

---

<sup>214</sup> Moore, *supra* note 11, at 1144.

<sup>215</sup> *Id.*

<sup>216</sup> *Id.*

<sup>217</sup> *Id.*

<sup>218</sup> CLAUS ROXIN, I DERECHO PENAL: PARTE GENERAL 807 (Diego-Manuel Luzón Peña et al. trans., Civitas 2d ed. 2000) (1994).

<sup>219</sup> *Id.* at 807–08.

<sup>220</sup> *Id.* at 808.

<sup>221</sup> *Id.*

<sup>222</sup> *Id.*

<sup>223</sup> *Id.* at 808–09.

argument shares several important features with Daniel Dennett's compatibilism. Jakobs, like Dennett, believes that the usefulness of certain concepts depends on their effectiveness in explaining, understanding, and predicting the behavior of a given system.<sup>224</sup> More specifically, Jakobs argues that free will is relevant to responsibility only insofar as it is practical and useful for law and society to presuppose that humans are free in a special sort of way.<sup>225</sup> Jakobs then posits that the freedom to be unencumbered from causal processes is not essential to our judgments about blame and responsibility.<sup>226</sup> However, he argues that being free from extraordinary causal influences, such as serious human threats and mental defects is essential to our practices of blaming and punishing.<sup>227</sup> Therefore, he suggests that judgments about moral responsibility are dependent on being free from certain compulsions, but not dependent on being free from causal processes.<sup>228</sup> As a result, Jakobs concludes that the sort of free will that is relevant to moral responsibility is not threatened by the truth of causal determinism.<sup>229</sup>

### C. Agnostic Stances Toward the Problem of Free Will in Criminal Law

In a recently published book, Larry Alexander seems to reject the compatibilist solution to the free will problem in favor of what seems to be an agnostic stance with regard to the issue. More specifically, he claims that "compatibilism provides only a hollow form of moral responsibility, not the full-blooded form that our reactive attitudes assume."<sup>230</sup> What Alexander finds particularly troubling about the compatibilism advocated by scholars like Morse and Dennett is that, if determinism obtains, it would seem that what appears to the actor to be a "reason" for engaging in a certain act is actually caused by forces that the actor cannot control.<sup>231</sup> He contends, "we are incapable of resolving" the "free

---

<sup>224</sup> Thus, for example, Dennett argues that there are good reasons to treat two chess programs that are competing against each other as "agents" that are capable of acting in different ways and of harboring "beliefs" about what the other will do. We have good reasons to do so because this is the best way of *understanding* what is going on when the two chess programs face each other. Whether chess programs have "true" beliefs or are "really" agents is beside the point. What matters is whether it is useful for us, in an explanatory sense, to treat them as agents that harbor beliefs about each other. See generally DENNETT, *supra* note 156, at 81.

<sup>225</sup> See generally Günther Jakobs, *Individuo y persona. Sobre la imputación jurídico-penal y los resultados de la moderna investigación neurológica*, in RAFAEL ALCÁ CER GUIRAO, *EL PROBLEMA DE LA LIBERTAD DE ACCIÓN EN EL DERECHO PENAL* 154 (2007) (stating that there are good reasons for social systems in general and legal systems in particular to assume that humans possess a certain kind of free will even if "true" free will is ruled out by determinism).

<sup>226</sup> *Id.* at 152.

<sup>227</sup> *Id.* at 152–53.

<sup>228</sup> *Id.*

<sup>229</sup> *Id.*

<sup>230</sup> ALEXANDER & FERZAN, *supra* note 35, at 15.

<sup>231</sup> *Id.*



will-determinism puzzle” and that the specter of determinism “will always dog [our] practices of holding people morally responsible.”<sup>232</sup> However, he believes that “we cannot imagine dispensing with” our practices of blaming and punishing.<sup>233</sup> Given that we cannot live without such practices, he concludes “a retributivist regarding criminal punishment need not resolve or even take sides on the free will issue.”<sup>234</sup> Alexander then summarizes his position by claiming in a somewhat mysterious fashion that “we cannot . . . comprehend the bases of moral responsibility” because “neither determinism nor indeterminism can provide a satisfactory account of moral responsibility, and together they appear to exhaust the possibilities.”<sup>235</sup> Nevertheless, he posits that we cannot renounce to judgments about blame and praise because “we cannot comprehend the possibility that we are *not* morally responsible.”<sup>236</sup>

In his book on the philosophy of criminal law, Douglas Husak also seems to adopt an agnostic stance to the free will problem. There, Husak argues that the gist of the criminal law’s act requirement cannot mean that an actor should only be held liable for engaging in “acts.”<sup>237</sup> The act requirement, Husak suggests, cannot stand for this proposition because actors are routinely held liable for non-acts such as omissions and for certain states of affairs such as public drunkenness.<sup>238</sup> Therefore, he claims that what the criminal law’s act requirement really means is that it is unfair to hold an actor liable for acts, results, or states of affair that he cannot control.<sup>239</sup> Husak thus contends that the criminal law should abandon the act requirement in favor of what he calls the “control principle.”<sup>240</sup> One obvious objection to Husak’s control principle is that if determinism obtains, an actor should never be held liable for anything, given that in a causally determined world human behavior is ultimately caused by factors over which the individual lacks control. To his credit, Husak acknowledges the problem in his book.<sup>241</sup> Nevertheless, he ends up dismissing the problem by asserting “though perhaps disappointingly little is said here” about how determinism might undermine the control principle, “I hope that the notion of control is able to withstand the great intuitive weight I place upon it.”<sup>242</sup>

---

<sup>232</sup> *Id.*

<sup>233</sup> *Id.*

<sup>234</sup> *Id.*

<sup>235</sup> *Id.* at 15 n.22.

<sup>236</sup> *Id.*

<sup>237</sup> HUSAK, *supra* note 32, at 97–98.

<sup>238</sup> *Id.* at 83–90.

<sup>239</sup> *Id.* at 98.

<sup>240</sup> *Id.*

<sup>241</sup> *Id.* at 98 (stating that “I do not dispute the relevance of the [free will problem] to criminal theory revised to include the control principle”).

<sup>242</sup> *Id.*

*D. Hard Incompatibilism and Criminal Theory*

For reasons that remain unclear, hard incompatibilism has barely found its way into the writings of contemporary criminal scholars. Only Michael Corrado has expressly endorsed the hard incompatibilist position, although he has yet to comprehensively defend his position against the attacks leveled by compatibilists and libertarians.<sup>243</sup> Additionally, a handful of criminal theorists defend a particular account of excuse defenses that is influenced by some of the arguments that hard incompatibilists often advance. This account is known as the “causal theory of excuses.”<sup>244</sup> According to the causal theory of excuses, a defendant ought to be excused for engaging in wrongful conduct if his conduct was caused by factors outside the actor’s control.<sup>245</sup> This view stems from two deeply held intuitions. First, it assumes that it is intuitively unfair to blame an actor for acts that are not “up to him.” Second, it assumes that it is unjust to blame an actor for performing a certain act if he lacked the capacity to do otherwise.

Supporters of the causal theory of excuses have a tough time accepting the implications that causal determinism would have for their theory. As has been mentioned, the truth of causal determinism means that human conduct is caused by the combination of natural laws and events that have taken place in the past.<sup>246</sup> Humans lack control over both natural laws and the past. Therefore, the truth of causal determinism inevitably leads to the conclusion that human conduct is the product of factors over which we lack control. Given that defenders of the causal theory of excuses believe that conduct caused by factors the actor cannot control should not be punished, the truth of causal determinism should lead them to reject free will and with it the practice of blaming people for their evil acts.<sup>247</sup> Causal theorists have resisted the implications of this argument. Most, like Anders Kaye, have done so by refusing to fully accept the thesis of causal determinism.<sup>248</sup> Kaye’s position is inspired, at least in part, by the plausibility of certain libertarian accounts of human behavior<sup>249</sup> and, therefore, his approach to determinism appears to be a brand of libertarianism cloaked in (partial) determinist clothing.

---

<sup>243</sup> See, e.g., Michael L. Corrado, *The Abolition of Punishment*, 35 SUFFOLK U. L. REV. 257, 265 n.20 (2001) (stating that “[i]f you accept the plausibility of determinism, then you should probably reject the theologically based notion of moral responsibility” and that “the idea of *moral* responsibility, upon which desert is based, requires genuine metaphysical freedom” and that “such freedom is inconsistent with the most plausible view of the world”).

<sup>244</sup> For an overview of the causal theory of the excuses, see Anders Kaye, *Resurrecting the Causal Theory of the Excuses*, 83 NEB. L. REV. 1116 (2005).

<sup>245</sup> *Id.* at 1120.

<sup>246</sup> See Fischer, *supra* note 117.

<sup>247</sup> This is the standard objection to the causal theory of excuses. See, e.g., MOORE, *supra* note 11, at 1112.

<sup>248</sup> Kaye defends what he calls “selective determinism” which, in turn, is a type of partial determinism. Kaye, *supra* note 244, at 1134.

<sup>249</sup> *Id.* at 1137.

V. THE FREE WILL MAZE AS A DIALECTICAL STALEMATE AND WHERE WE GO FROM HERE

As we have seen, philosophers and criminal theorists have come up with different ways of tackling the free will problem. Both libertarians and hard incompatibilists believe that causal determinism is incompatible with free will and moral responsibility. Libertarians, however, believe that humans do in fact have free will because human behavior can be traced back to indeterministic processes. Hard incompatibilists, on the other hand, argue that humans lack free will and moral responsibility even if human conduct is the product of indeterministic forces. Compatibilists do not deny that causal determinism is plausible. However, unlike hard incompatibilists and libertarians, they argue that free will is compatible with determinism. Although each of these views has some appealing attributes, important objections can be leveled against all three. The purpose of this Part is to briefly examine various objections to each of these theories in order to assess the current state of the free will debate and propose a way forward.

A. *The Problems with Libertarianism*

The most sophisticated account of libertarianism is the one defended by Anders Kane who, in typical libertarian fashion, accepts that free will is incompatible with determinism, but rejects the truth of causal determinism.<sup>250</sup> The challenge for Kane is to find a scientifically plausible account of human behavior that can be traced back to indeterministic processes.<sup>251</sup> While Kane's Herculean attempt to explain human conduct by appealing to indeterministic quantum mechanic processes in the brain<sup>252</sup> is creative, it raises two serious problems. First, the indeterministic processes that undergird quantum mechanics are random.<sup>253</sup> This is problematic, for it would be odd to conclude that it is fair to hold an individual morally responsible for conduct that is ultimately the product of random or haphazard events. Thus, random causation is as much a threat to freedom of will and moral responsibility as determinism is.<sup>254</sup> Second, and perhaps more importantly, there is simply no scientific evidence supporting Kane's account of quantum mechanics in the brain. As Henrik Walter observes, "to date there is no

---

<sup>250</sup> See *supra* notes 136–142 and accompanying text.

<sup>251</sup> This is what Kane calls the "The Intelligibility Problem." Kane, *supra* note 134, at 9. The challenge posed by "The Intelligibility Problem" requires the libertarian to "show that a libertarian free will requiring indeterminism can be made intelligible and how, if at all, such a free will can be reconciled with modern scientific views of the cosmos and human beings." *Id.*

<sup>252</sup> See *supra* notes 136–142 and accompanying text.

<sup>253</sup> Kane is aware of this problem. Thus, he acknowledges that "events that are undetermined, such as quantum jumps in atoms . . . happen merely by chance" and that it is difficult to understand how chance events can ground free and responsible conduct. Kane, *supra* note 134, at 9.

<sup>254</sup> Pereboom, *supra* note 30, at 485.

solid empirical evidence that local quantum phenomena play a role in neurons, [whereas] there are good arguments to the contrary.”<sup>255</sup>

The most recent criminal law theorist to defend libertarianism is George Fletcher. Fletcher believes that Noam Chomsky’s showing that we command an infinite number of sentences in natural language demonstrates that human conduct is not causally determined.<sup>256</sup> More specifically, he argues that it would be impossible to program a computer to create terms the programmer has never heard before and, therefore, it is unlikely that human beings could be determined to invent new words in a language.<sup>257</sup> However, Fletcher’s use of Chomsky as his poster boy for indeterminism is particularly odd, because other philosophers invoke Chomsky’s theory of language as an argument in favor of the so-called *computational theory of mind*.<sup>258</sup>

According to the computational theory of mind, the best way of thinking about the human mind is as an information processing system (i.e., a computer)<sup>259</sup> and the best way of thinking of the concept of thought is as a kind of computation.<sup>260</sup> Chomskyan linguistics lends credence to the computational theory of mind because what accounts for the infinite variability of language that Fletcher finds so interesting is the human capacity for “recursion.”<sup>261</sup> Recursion is nothing more than a particular type of computation that consists of taking discrete elements (e.g., phrases) and recombining them in a way that can potentially yield infinite combinations (e.g., infinite sentences).<sup>262</sup> Of course, given that recursion is a kind of computation, there would be nothing odd about a sufficiently powerful computer that can recombine a discrete set of words and phrases in a way that produces infinite utterances. Therefore, rather than demonstrating that human language cannot be the product of computation or of some algorithmic program, Chomskyan linguistics tend to show that the human capacity for language can be reduced to a discrete set of rules that operate in recursive fashion to create an infinite amount of utterances. As a result, contrary to what Fletcher suggests,

---

<sup>255</sup> HENRIK WALTER, NEUROPHILOSOPHY OF FREE WILL: FROM LIBERTARIAN ILLUSIONS TO A CONCEPT OF NATURAL ANATOMY 162 (2001), *cited in* Manuel Vargas, *Revisionism*, in *FOUR VIEWS ON FREE WILL* 144.

<sup>256</sup> Fletcher, *supra* note 22, at 278.

<sup>257</sup> *Id.* at 279.

<sup>258</sup> *See, e.g.*, Charles R. Gallistel, *Learning Organs*, in JEAN BRICMONT & JULIE FRANCK, *CHOMSKY NOTEBOOK* 197 (2009) (stating that “that the brain’s behaviorally relevant activity can be understood in computational terms is the computational theory of mind, which is at the core of contemporary cognitive science” and that “Chomsky’s work did much to establish this conceptual framework, in which the brain’s activity is conceived of in computational terms”)

<sup>259</sup> *Id.*

<sup>260</sup> JOHN-MICHAEL KUCZYNSKI, *CONCEPTUAL ATOMISM AND THE COMPUTATIONAL THEORY OF MIND* 10 (2007) (stating that, according to the computational theory of mind, “thinking is ‘computing’”).

<sup>261</sup> VIVIAN JAMES COOK & MARK NEWSON, *CHOMSKY’S UNIVERSAL GRAMMAR: AN INTRODUCTION* 18 (2d ed. 1997).

<sup>262</sup> HARRY VAN DER HULST, *RECURSION AND HUMAN LANGUAGE*, at xv–xvi (2010).

Chomsky's theory of language reinforces the thesis of causal determinism instead of jeopardizing it.

Fletcher also believes that John Searle's Chinese Room argument shows that human conduct is likely not causally determined.<sup>263</sup> Fletcher argues that "if the use of language were determined by algorithms in the brain," we would conduct ourselves in much the same manner as the individual inside the Chinese Room and would thus be unable to understand what we're saying.<sup>264</sup> Since we clearly understand what we're saying, it follows that we are not determined to act by algorithms in the brain. The chief objection to Fletcher's use of the Chinese Room argument as a vehicle for demonstrating the falsity of determinism is that the Searle's argument is not designed to demonstrate that human conduct is not causally determined. Searle devised the argument in order to show that computers, unlike humans, are not capable of understanding meaning.<sup>265</sup> Even if this conclusion is accepted,<sup>266</sup> it does not follow from the fact that humans are not computers that human conduct is not causally determined. The basic flaw in Fletcher's argument is that it incorrectly assumes that all causally determined beings take the form of a computer. Thus, as Stephen Morse correctly points out, Fletcher's argument ultimately fails because "[c]omputers are not the only deterministic mechanisms on earth, and Fletcher fails to confront the possibility that computer programs cannot fully simulate the capacity of natural selection to develop new, nonmysterious and deterministic forms, such as brains capable of understanding meaning."<sup>267</sup>

### B. *The Problems with Compatibilism*

#### 1. *Basic Objections to Standard Accounts of Compatibilism Referred to as "Mesh" and "Reactive Attitudes" Compatibilism*

Frankfurt claims that the sort of free will that undergirds judgments of blame and praise merely requires that there be an appropriate "mesh" or link between certain inclinations and desires of the actor.<sup>268</sup> More specifically, Frankfurt suggests that an actor acts freely if his first and second order volitions are connected in the right kind of way (i.e., the actor desires to desire eating the red velvet cheesecake).<sup>269</sup> This solution is elegant because it allows for the possibility of free will and moral responsibility even if it turns out that we are causally determined to act in a certain manner (to eat the red velvet cheesecake, for

---

<sup>263</sup> Fletcher, *supra* note 22, at 279; *see also supra* notes 194–195 and accompanying text.

<sup>264</sup> *Id.*

<sup>265</sup> *See supra* notes 192–197 and accompanying text.

<sup>266</sup> Many philosophers reject Searle's conclusion. *See, e.g.,* DANIEL C. DENNETT, *THE INTENTIONAL STANCE* 323–26 (1996).

<sup>267</sup> Morse, *supra* note 34, at 2550.

<sup>268</sup> *See supra* notes 163–169 and accompanying text.

<sup>269</sup> *See supra* notes 161–167 and accompanying text.

example). The problem with Frankfurt's solution is that causal determinism threatens not only the control humans have over their actions, but also the control they have over their motives, intents, and desires.<sup>270</sup> Therefore, if causal determinism is true, it follows that the interplay of the past and natural laws determines not only our conduct, but also our desires. Given that we lack control over the past and the natural laws, it follows that if determinism obtains we would lack control over both our conduct *and* our *first, second, and Nth order volitions*. Once we accept that in a causally determined universe humans lack control over both their conduct and their volitions, it is difficult to see why the actor's volitions make him responsible in a robust sense for what he does. As a result, mesh accounts of compatibilism, like Frankfurt's, do not seem to provide the sort of free will that supports judgments of moral responsibility. Since an actor's responsibility seems to be a product not only of his capacity to act otherwise, but also of his capacity to desire otherwise—both of these capacities are threatened by determinism.

P.F. Strawson attempts to avoid objections to mesh theories of compatibilism with this proposition: we must assume that humans are endowed with free will, regardless of the truth of determinism, because not doing so would lead to abandoning many reactive attitudes (regret, resentment, blame, praise, love, etc.) that are essential to healthy interpersonal relationships.<sup>271</sup> Strawson's argument is thus more normative than metaphysical.<sup>272</sup> Ultimately, his claim is not that we actually have such a thing as free will, but rather that human and societal life is more *appealing* if we assume we have free will regardless of whether we actually do have it.<sup>273</sup> Certainly, Strawson is right that some reactive attitudes, such as the kind of blame that undergirds theories of retribution, would have to be jettisoned if we lack free will. However, it is unclear whether reactive attitudes are actually essential to interpersonal relationships. Even if they are, it is unclear whether assuming that we lack free will would jeopardize the reactive attitudes that are most essential to human life. As will be discussed in Part VI of this article, we

---

<sup>270</sup> Philosophers call this the "ultimate responsibility" ("UR") objection. See ROBERT KANE, *THE SIGNIFICANCE OF FREE WILL* 73 (1996). According to UR, genuinely responsible actors "must not only be the sources of their actions, but also of the *will* to perform the actions." *Id.* As a result, UR would hold that an actor is not responsible for her acts merely because she desired to desire what she did. These actors would only be free if they are the sources of their first and second order volitions. If, on the other hand, actors are not the ultimate sources of their first, second, and Nth order volitions, they are not responsible for their acts. Causal determinism seems to imply that everything that takes place in the universe, *including human desires and motivations* is the product of the convergence of the past and natural laws. It would thus seem that what produces human volitions is something beyond the control of the actor. Therefore, UR would lead to denying responsibility in Frankfurt type cases.

<sup>271</sup> See *supra* notes 170–176 and accompanying text.

<sup>272</sup> Pereboom, *supra* note 30, at 148 (stating that Strawson mounts a normative defense of reactive attitudes that should hold even if causal determinism obtains).

<sup>273</sup> Strawson, *supra* note 18, at 55–56.

experience many of the most cherished reactive attitudes, including love, admiration, and praise, regardless of whether we believe that the actor who is loved, praised, or admired is endowed with free will.<sup>274</sup> Thus, most of us would continue loving our children, admiring Liz Taylor's beautiful violet eyes, and praising Ussain Bolt for running a hundred meters in less than 9.69 seconds, even if we believed that our children have not yet developed the sort of character or rationality that undergirds free will, that Liz Taylor did not freely choose to have violet eyes, and that Ussain Bolt's freakish speed is more a product of his genes than his training. It is thus uncertain that Strawson is right when he claims that a life without free will is not possible or desirable.

## 2. *The Problems with Reasons-Centered Accounts of Compatibilism (and Semi-Compatibilism)*

As we have seen, semicompatibilists like John Martin Fischer believe that determinism is incompatible with free will, while it is compatible with moral responsibility.<sup>275</sup> For Fischer, an actor is morally responsible if his conduct could be guided by reasons, and the reasons that move him into action belong to him rather than to another.<sup>276</sup> The flaws with Fischer's account of moral responsibility are similar to the flaws inherent in Frankfurt's mesh account of compatibilism. If causal determinism is true, it follows that everything that happens in the world, including human conduct, human desires, and a human's responsiveness to reasons is caused by factors over which he lacks control.<sup>277</sup> It is unclear whether an actor should be blamed or praised for engaging in conduct that is the product of a reason-responsive process of deliberation when it is assumed that the very process of deliberation was determined by factors over which the actor lacked control. It is sensible to argue that an actor is morally responsible for his acts if, and only if, the process of deliberation that caused his conduct originated in the actor rather than in some causal force over which the actor has no control.<sup>278</sup> In other words, it is reasonable to assume that an actor is morally responsible for his conduct only if he was in control of the deliberative process that caused him to act. Determinism implies that actors do *not* control the deliberative processes that generate their conduct. Thus, it may be argued that the question of whether the actor is in fact responsive to reasons in a particular case lacks the moral significance that Fischer attaches to it, because the actor does not control the deliberative process.

Like Fischer, compatibilist criminal theorists believe that if an actor has the capacity for rationality, then he also has the sort of free will that undergirds judgments of blame and praise.<sup>279</sup> These arguments can be criticized on the same

---

<sup>274</sup> Pereboom, *supra* note 30, at 483–87 (defending this position forcefully).

<sup>275</sup> Fischer, *supra* note 117, at 78–79.

<sup>276</sup> *Id.* at 78–79.

<sup>277</sup> *Id.* at 73–80.

<sup>278</sup> KANE, *supra* note 270 (discussing the “Ultimate Responsibility” argument).

<sup>279</sup> Morse, *supra* note 203, at 441.

grounds that are employed on Fischer's semicompatibilism arguments.<sup>280</sup> Moore and Morse, however, attempt to sidestep these objections by deploying a *reductio ad absurdum* argument. They argue that if humans can only be blamed for conduct that is the product of factors that they control, then it follows that humans can never be blamed for anything, given that causal determinism implies that all human conduct is the product of forces that we cannot control.<sup>281</sup> And they claim it would be unpalatable to conclude that everyone ought to be excused for their transgressions.<sup>282</sup> The problem with this sort of maneuver is that, like Strawson's reactive attitudes compatibilism, it presupposes that if we lack the capacity to be held morally responsible for our acts, this leads to an impoverished view of societal life that would "falsify much of our moral li[ves.]"<sup>283</sup> Moore and Morse (and Strawson) overstate their case, because (as argued previously) it is not clear that assuming a lack of free will would lead to the sort of life that is not worth living. Furthermore, accepting that everyone ought to be "excused" from wrongdoing in a *retributive* sense does not entail that no one ought to be "punished" or "incarcerated." As will be discussed in Part VII, there are good consequentialist reasons to punish or incarcerate people even if they do not "deserve" to suffer in the sense that Morse and Moore imply.

### 3. *The Problems with Pragmatic or Functionalist Accounts of Compatibilism*

Daniel Dennett and Gunther Jakobs argue that it is useful to talk about concepts like free will only if doing so helps us better understand and explain the behavior of a particular being or entity.<sup>284</sup> Thus, we ought to abstain from talking about the free will of ants or rocks because there is nothing to gain in terms of explanatory and predictive power. However, it is perfectly sensible and rational to talk about the free will of humans, because doing so helps us to better understand societal practices of blaming and punishing and the way in which humans interact with each other.<sup>285</sup> This approach to free will and moral responsibility views the problem of free will as a pragmatic one that should not be obfuscated by asking unanswerable metaphysical questions, such as whether humans actually have free will. Ultimately, it is unimportant whether humans have metaphysical free will; what really matters is whether there is something to gain by assuming that they are, in fact, endowed with free will.<sup>286</sup>

---

<sup>280</sup> ALEXANDER & FERZAN, *supra* note 35, at 15 (stating that Larry Alexander believes that reason-centered compatibilism "seems unresponsive to the worry that what appears to an actor to be a reason, or a reason with a particular positive or negative weight, seems to be beyond the actor's proximate control").

<sup>281</sup> Moore, *supra* note 11, at 1112–13.

<sup>282</sup> *Id.* at 1113.

<sup>283</sup> *Id.* at 1145.

<sup>284</sup> See *supra* notes 224–229 and accompanying text.

<sup>285</sup> See DENNETT, *supra* note 156, at 81.

<sup>286</sup> DANIEL C. DENNETT, *ELBOW ROOM* 163–64 (1984) (stating that whether we assume that we have free will and are capable of moral responsibility should not be



The problem with this argument is that many people do in fact care about whether we actually have the ability to control our conduct and the desires and reasons that shape our behavior. For those who do care about this sort of thing, determinism threatens to undermine free will and moral responsibility. Whether it would be efficient or pragmatic to ignore this is irrelevant. While there may be something to be gained from adopting a pragmatic approach to questions related to free will, those who believe that judgments of blame are unjust if humans do not have actual access to alternate possibilities will not change their minds simply because it would be efficient to presume that they do have the capacity to choose amongst alternative courses of action.

### *C. The Problems with Hard Incompatibilism*

Many hard incompatibilists believe both that determinism is true and that it is incompatible with free will and moral responsibility. This kind of hard incompatibilism can be challenged on two fronts. First, it is unclear whether causal determinism fully explains human behavior. It is true that there are scientific experiments that suggest that human conduct is causally determined by factors over which we lack control,<sup>287</sup> but it is also true that these experiments are limited in their scope and the precise way in which the mind works is still very much unknown.<sup>288</sup> Furthermore, even if it is accepted for argument's sake that determinism is true, it is possible that the sort of freedom that is essential for moral responsibility is compatible with a deterministic understanding of the universe and of human behavior. If there are, as many philosophers and legal scholars seem to believe, accounts of human conduct that are compatible both with determinism and with ascriptions of blame and praise, why hijack the free will edifice upon which our practices of blaming and punishing are built, simply because it is not incoherent to hold that determinism is incompatible with free will and moral responsibility? After all, as Stephen Morse suggests, we may have good reason to reject hard incompatibilism “[if] compatibilism is consistent with our responsibility practices and their centrality, and [if] there is [not] and cannot be any incontrovertible theoretical or empirical reason to reject it . . . .”<sup>289</sup> Thus, although “there may be good empirical and normative reasons to reform various responsibility doctrines and practices . . . there is no metaphysical reason concerning free will to abandon them entirely.”<sup>290</sup>

---

dependent on a metaphysical question, but rather ought to depend on whether assuming freedom and moral responsibility ties well with some “recognizable social desideratum”).

<sup>287</sup> See, e.g., Libet, *supra* note 8, at 640.

<sup>288</sup> DENNETT, *supra* note 156, at 227–42 (criticizing experiments such as Libet’s).

<sup>289</sup> Morse, *supra* note 34, at 2553.

<sup>290</sup> *Id.*

*D. The Free Will Problem as a Dialectical Stalemate*

This brief survey of the proposed solutions to the free will problem—and the objections that can be leveled at the different solutions—demonstrates at least three things. First, causal determinism threatens to undermine free will and proponents of free will (libertarians) thus need to put forth a theory that can explain either why free will is compatible with determinism or why determinism is likely false. Second, some theorists (compatibilists) have in fact come up with plausible theories that purport to demonstrate how free will and determinism can coexist, whereas others (libertarians) have come up with plausible theories that explain why determinism might be false. Third, serious objections can be directed both at the incompatibilist (including libertarian and hard incompatibilist theories) and compatibilist solutions to the free will problem. As a result, it seems that the free will problem has the argumentative structure John Martin Fischer called a “dialectical stalemate.”<sup>291</sup>

The seeds of a dialectical stalemate are planted whenever someone puts forth a controversial claim and supports his argument by “invoking a set of examples (or other considerations)” that demonstrate that his claim ought to be accepted.<sup>292</sup> The dialectical stalemate germinates when an opponent reasonably asserts that one could “embrace all the examples” (and other considerations) in the proponent’s argument without having to accept the proponent’s claim.<sup>293</sup> Thus, as Hillary Bok explains, we have reached a dialectical stalemate

when no arguments based on appeals to ordinary language, to the consideration of examples, or to our intuitions succeed in convincing our opponents, and when this is due not to our opponents’ limitations but to the fact that both sides can appeal to intuitions, accounts of apparent counterexamples, and claims about our ordinary use of the terms in question that are not unreasonable.<sup>294</sup>

Once the notion of a dialectical stalemate is grasped, it is easy to see why the free will problem presents the “signature structure” of such argumentative deadlocks.<sup>295</sup> Libertarians and hard incompatibilists share the intuition that an actor lacks the free will that calls for moral responsibility, and the intuition that determinism implies that humans lack control over their actions. Compatibilists reply in one of two ways. Some offer the competing intuition that in many cases moral responsibility is grounded on something other than the capacity to control

---

<sup>291</sup> FISCHER, *supra* note 38, at 112–17.

<sup>292</sup> *Id.* at 112–13.

<sup>293</sup> *Id.*

<sup>294</sup> HILLARY BOK, *FREEDOM AND RESPONSIBILITY* 99 (1998).

<sup>295</sup> FISCHER, *supra* note 38, at 116.

actions, such as the appropriate connection between the actor's first and second order volitions.<sup>296</sup>

Others argue that even if determinism obtains, humans can act otherwise—for any given actor could act otherwise if he wants to, because even if determinism is true, an actor's desires do exert causal force over the actor's conduct.<sup>297</sup> Therefore, as Hillary Bok points out,

both [incompatibilists] and compatibilists draw on important features of our ordinary concept of freedom, both are trying to apply that concept to cases in which the conditions of its straightforward application are absent, both project that concept in ways that are not obviously unreasonable or illegitimate, and both can muster real intuitive support for their views.<sup>298</sup>

Consequently, the free will problem presents a genuine dialectical stalemate, given that “no appeal to our ordinary concept of freedom, or to the ways in which we ordinarily apply it, will settle the issue between [incompatibilists] and compatibilists, since that concept supports both views and does not give us decisive grounds to reject either.”<sup>299</sup> More importantly, if the problem of free will truly generates this dialectical stalemate, incompatibilists and compatibilists should agree that both solutions to the free will problem “have something to be said for them [and] can usefully be employed in various circumstances.”<sup>300</sup> They should also agree that neither claim is likely “to be established by arguments about what [ordinary terms mean], by appeal to our ordinary concept of freedom or by our intuitions.”<sup>301</sup>

#### *E. The Way Out of the Dialectical Stalemate Generated by the Free Will Problem*

What should criminal scholars do in light of this dialectical stalemate? We could throw our hands up in despair and candidly acknowledge, as Larry Alexander has, that the free will problem is intractable and one that “we are incapable of resolving.”<sup>302</sup> This strikes me as the wrong way to respond to a dialectical stalemate. As John Martin Fischer has persuasively argued, these stalemates should not issue in “philosophical despair,” nor should they “result in our inability to make any philosophical progress or to come to any useful

---

<sup>296</sup> FRANKFURT, *supra* note 115.

<sup>297</sup> For a survey of the different attempts to justify compatibilist views of free will on the basis of this conception of what it means to be able to act otherwise, see Bernard Berofsky, *Ifs, Cans and Free Will: The Issues*, in THE OXFORD HANDBOOK OF FREE WILL 181, 188–95.

<sup>298</sup> BOK, *supra* note 294, at 99.

<sup>299</sup> *Id.*

<sup>300</sup> *Id.*

<sup>301</sup> *Id.* at 100.

<sup>302</sup> ALEXANDER & FERZAN, *supra* note 35, at 15.

philosophical conclusions.”<sup>303</sup> When faced with these stalemates, we should “abandon the attempt to convince our opponents that intuitions, examples, or ordinary language decisively favor our view, not only because such arguments are unlikely to convince them[,] but because the fact that we have reached a genuine dialectical stalemate shows that such arguments are unsound.”<sup>304</sup> And we *should* “admit that there are several apparently legitimate ways” of solving the problem and that each of these solutions is *prima facie* plausible.<sup>305</sup> Once we do so, we ought to stop focusing on whether free will *really* or *actually* means this or that, or whether moral responsibility *does* or *does not* presuppose access to alternative possibilities. Rather, we should ask ourselves whether we have better *normative* reasons for adopting the compatibilist over the incompatibilist stance or vice versa.

In other words, the best way out of the dialectical stalemate is not by considering which of the competing accounts of free will is closer to the metaphysical truth of the matter (because we simply do not know what the truth of this matter is), but rather by considering the *normative* question related to which of these competing conceptions of freedom and moral responsibility *produces a more appealing life in general and a more desirable criminal law in particular*.

#### VI. WHY LIVING WITHOUT FREE WILL IS NOT AS BAD AS IT SEEMS

Before deciding whether there are good normative reasons for us to hold on to—or abandon—free will, we first need to envision what a world without free will would look like. Would we want to live in such a world or would we rather live in a world in which we assume that others are endowed with free will? For most of the philosophers and criminal theorists who have confronted the question, the answer seems obvious. A life without free will is like a garden without flowers. Assuming that our fellow humans lack free will and the capacity to be held morally responsible for their acts leads to an impoverished view of human life.<sup>306</sup> The purpose of this Part is to show that this conventional account of what a world without free will would look like is wrong in very significant ways. It seems that these bleak assessments of the consequences of assuming that we have no free will are driven by the philosopher’s desperate desire to hang on to the free will edifice that we have built over thousands of years. Ultimately, however, this Part claims that the normative case against free will is overstated, that there is no need to fear living in a world without free will, and that assuming that humans lack free will can actually have salutary consequences in some cases.

---

<sup>303</sup> FISCHER, *supra* note 38, at 116.

<sup>304</sup> BOK, *supra* note 294, at 99.

<sup>305</sup> *Id.* at 99–100.

<sup>306</sup> This is the view espoused by P.F. Strawson and Michael Moore. *See* Strawson, *supra* note 18; Moore, *supra* note 11, at 1142 n.138.

A. *The Conventional Claim: Living Without Free Will Is Utterly Unappealing*

1. *Strawson's Reactive Attitudes Argument*

P.F. Strawson argues that we ought to assume that humans are endowed with free will because to do the opposite undermines the reactive attitudes.<sup>307</sup> As discussed earlier, reactive attitudes are attitudes that are important in defining and maintaining our interpersonal and societal relationships.<sup>308</sup> For Strawson, a life lived without reactive attitudes is barely a life worth living.<sup>309</sup> After all, what would a life be without receiving praise for doing well and blaming those who do badly? What would a life be without being able to love those who are good to us and resent those who are mean to us? The answer is clear to Strawson: living such a life is unpalatable and ought to be avoided at all costs.<sup>310</sup> If the price of not falling prey to such a life is to assume that we have free will in the face of causal determinism, so be it.

Strawson believes that assuming that we do not have free will jeopardizes not only the enjoyment of individual lives, but also the wellbeing of society as a whole.<sup>311</sup> Reactive attitudes are thought to be essential to justifying many societal practices. Perhaps the most obvious practices that are justified on the basis of certain reactive attitudes are practices of blaming and punishing.<sup>312</sup> Religious practices also seem to be undergirded by certain reactive attitudes. Judgments of blame and praise play an essential role in Judeo-Christian religions. For example, for the Judeo-Christian tradition, a person who unjustifiably violates one of the Ten Commandments behaves in a blameworthy manner. Desert-based judgments are thus essential to understanding certain basic aspects of these religions. Furthermore, even more mundane practices, such as governmental awarding of “merit scholarships” and public recognition of achievements by way of prizes and awards, seem to lose meaning in a world without free will. After all, if there is no free will, what’s the point of rewarding people for doing what they could not have abstained from doing? By the same token, if there is no free will, what’s the point of blaming people for doing things that they could not have done differently? Strawson argues that these practices would be pointless if we assume that there is no free will and this, in turn, provides us with a normative argument against hard incompatibilist approaches to the free will problem.<sup>313</sup>

---

<sup>307</sup> Strawson, *supra* note 18.

<sup>308</sup> *Id.* at 47–50.

<sup>309</sup> *See id.* at 55–56.

<sup>310</sup> *See id.*

<sup>311</sup> *See id.* at 65 (“[A]n awareness of variety of forms should not prevent us from acknowledging also that in the absence of *any* forms of these attitudes it is doubtful whether *we* should have anything that we could find intelligible as a system of human relationships, as human society.”).

<sup>312</sup> *See id.* at 48–51.

<sup>313</sup> *See id.* at 62–66.

## 2. Michael Moore's "Falsifying Our Moral Life" Argument

Michael Moore has also assessed the normative implications of assuming that humans lack free will.<sup>314</sup> For him, the problem with denying free will is that doing so is inconsistent with our practices of blaming and punishing, praising, and rewarding. Like Strawson, Moore claims that our practices of blaming and praising are essential to moral life.<sup>315</sup> Without assuming that we are endowed with the sort of freedom that undergirds such practices we could not do things that we would like to do, such as praising (or blaming) Andy Warhol for immortalizing Campbell Soup cans in a work of art or commending Michael Moore for writing an excellent article about free will. Moore finds it hard to imagine what our practices of praising and blaming would look like if we assume that humans lack free will, although it seems obvious that he believes that such practices would be much less appealing if we were to make this assumption. While Moore concedes that the hard incompatibilist might adopt a "tough it out" attitude and claim that most of our moral experience is false in light of the truth of causal determinism, he believes that we have good reasons to avoid doing this.<sup>316</sup> Moore proposes that the hard incompatibilist position ought to be rejected, not because it can be proven false, but rather because it cannot be proven true and he asks if the implications of accepting the position would "falsify much of our moral life."<sup>317</sup> As a result, Moore concludes that we ought to reject hard incompatibilism and accept compatibilism.

## 3. The Attitudes that Must Be Abandoned in a World Without Free Will Are Not Essential to Maintaining Healthy Relationships

Strawson is probably right when he claims that some attitudes would have to be abandoned in a world without free will. He is also right when he claims, along with Moore, that it would no longer make sense to genuinely blame those who engage in wrongdoing and praise those who engage in morally commendable behavior. Nevertheless, Strawson overstates his case when he suggests that living without such attitudes is unpalatable or that the absence of such reactive attitudes is detrimental to societal life. Similarly, Moore exaggerates when he proposes that the loss of certain attitudes, such as blame, would "falsify much of our moral li[ves]."<sup>318</sup> While Strawson and Moore argue that feelings of blame are healthy both at individual and collective scales, there are good reasons to believe otherwise. Blame usually engenders resentment and indignation, which, in turn, might generate a strong desire for vengeance and an unrelenting quest for

---

<sup>314</sup> Moore, *supra* note 11, at 1145.

<sup>315</sup> *Id.* at 1144–45, 1148–49.

<sup>316</sup> *Id.* at 1145.

<sup>317</sup> *Id.*

<sup>318</sup> *Id.*

revenge.<sup>319</sup> While perhaps feelings of blame are not objectionable per se, there are many good reasons to object to vengeance and revenge. Vengeance and revenge are viewed as “angry” responses to crime that are derived more from a “lust for blood” than from a rational and careful consideration of the pros and cons of punishing the individual.<sup>320</sup> Thus, abandoning feelings of blame might lead to a less vengeful and violent society, which, of course, would be a salutary development.

It is also important to note that many of our most fulfilling relationships are entirely devoid of feelings of blame. Most people love to interact with children despite the fact that they cannot be genuinely blamed or praised for doing what they do.<sup>321</sup> Not only that, but it would seem that we prize such interactions at least in part precisely because we can easily shrug off transgressions committed by children by simply assuming that they could not do otherwise. Far from reducing the value of our experiences with children, this actually allows us to enjoy our time with them more, for we feel free to love children without judging them.

Similarly, most people cherish their pets although animals are clearly incapable of free will and thus of being genuinely responsible for their acts. In spite of this, we love spending time with our dogs and cats and very much appreciate their displays of affection and disregard their transgressions *without caring about whether they can be genuinely praised for curling next to us or blamed for biting our leg*. Thus, it is undoubtedly the case that experiencing “genuine” feelings of blame and praise is not a prerequisite to having fulfilling and meaningful relationships with others. When this is combined with the fact that feelings of blame can and often do degenerate in unhealthy desires for vengeance and revenge, Strawson and Moore’s conclusion that blame is essential to a life worth living seems far from compelling.

There are, of course, certain attitudes that could not be abandoned without dealing a fatal blow to some of the most cherished experiences in human life. The obvious example is love. It is difficult to imagine what a world not capable of loving would look like. Strawson seems to suggest that even our feelings of love would be threatened if we were to assume that we lack free will.<sup>322</sup> Once again, it seems that Strawson overstates his case. As was discussed in the preceding

---

<sup>319</sup> Blame relates to what Pereboom calls “moral anger.” Pereboom, *supra* note 30, at 488. This is the type of anger “that is directed toward someone who is believed to have done wrong.” *Id.* Pereboom cogently argues that although moral anger is sometimes useful, it frequently “fail[s] to contribute to the well-being of those against whom [it is] directed” and that moral anger “tends to damage or destroy relationships.” *Id.*

<sup>320</sup> JOEL FEINBERG, *DOING AND DESERVING: ESSAYS IN THE THEORY OF RESPONSIBILITY* 69 n.16, 103 (1970).

<sup>321</sup> *See id.* at 486.

<sup>322</sup> Strawson asks whether assuming that we lack free will would threaten the most basic reactive attitudes, including “all reciprocated adult loves.” Strawson, *supra* note 18, at 53.

paragraph, we love children and pets although we accept that they lack free will.<sup>323</sup> Furthermore, we would continue to love our children even if scientists demonstrate beyond a reasonable doubt that our love for them is entirely hardwired into our brains. By the same token, we would continue to love our parents, spouses, and friends even if someone showed us that we did not freely choose to love them or that they did not they freely choose to love us in return.<sup>324</sup>

Feelings of admiration are also very important to the human experience. It would be a bleak life indeed if no one would ever acknowledge our accomplishments. Moore believes that just like there is no room for blame, there is no space for praise in a world without free will.<sup>325</sup> Although it is unclear whether we must stop praising people if we assume that we lack free will, it is manifestly clear that we need not stop admiring them. We admired Liz Taylor's violet eyes—and she undoubtedly enjoyed such admiration—although we were aware that she did not freely choose the coloration of her iris. Similarly, we admire the Grammy award winning singer Adele's incredible voice even though we know that her vocal abilities are innate. Furthermore, we would continue to admire Ussain Bolt for being the fastest person on earth and to complement Michael Moore for his enormous contributions to legal theory even if we believed that they are able to do what they do only because they were lucky to have been born with certain physical and intellectual attributes. In sum, it appears that we often seek admiration and that we can admire humans for their traits and acts regardless of whether they have the ability to control their traits or acts.

#### *4. The Existence and Usefulness of Morality Is Not Threatened by Assuming that We Lack Free Will*

Another argument against the hard incompatibilist position is that its adoption might lead to eschewing moral judgments. This is an important objection. Moral rules are a fundamental feature of societal life. A theory that leads to the rejection of morality thus fails to make sense of an essential facet of our lives. This argument against hard incompatibilism exploits a perceived connection between judgments about blame and judgments about morality. The point of departure of the argument is that if humans lack free will they cannot be genuinely blamed for their acts. The next step is to argue that judgments about the morality of engaging in certain acts are devoid of meaning if one accepts the proposition that ascriptions of blame for engaging in these acts are unwarranted. After all, if it does not make sense to blame an actor for doing "X," is it not also the case that it does not make sense to conclude that the actor behaved immorally when doing "X"?

Despite its intuitive appeal, this argument ought to be rejected because it fails to distinguish between two very different rules. More specifically, the argument

---

<sup>323</sup> For the purpose of this article, I have assumed that children and non-human animals lack free will.

<sup>324</sup> Pereboom, *supra* note 30, at 486.

<sup>325</sup> Moore, *supra* note 11, at 1144–45.



conflates rules that govern the attribution of culpability with rules that proscribe wrongful conduct. Failing to distinguish between these two kinds of rules is problematic, given that it is possible to have wrongdoing without culpability. An act is *wrongful* if it is prohibited by a given system of norms (moral, legal, religious, and so forth).<sup>326</sup> On the other hand, an actor has *culpability* if he deserves blame for engaging in the admittedly wrongful conduct.<sup>327</sup> The foundational rules of morality (do not kill, do not steal, and so forth) are rules that proscribe wrongdoing rather than rules that gauge the actor's culpability. Judgments about the morality of engaging in certain conduct are thus independent from judgments about whether an actor should be blamed for engaging in an immoral act.<sup>328</sup> A case in which judgments of wrongdoing are divorced from judgments of blame is that of insane killings, where killing an innocent human being is deemed wrongful even where the actor is ultimately excused because he cannot be fairly blamed for the killing.<sup>329</sup>

The distinction between wrongdoing and blameworthiness also helps to make sense of certain important features of tort law. As any first-year law student knows, children and the insane are held liable for their torts even though their conduct is not usually considered blameworthy for the purposes of the criminal law.<sup>330</sup> Therefore, as far as tort law is concerned, the wrongfulness of certain conduct is determined solely by the fact that the act unjustifiably infringes a rule of the system regardless of whether the infraction is blameworthy.<sup>331</sup> Harming an innocent human being is morally and legally wrongful, even when the person who causes the harm is a minor or is insane. Once the independent moral significance of wrongdoing is grasped, it becomes clear why acceptance of the hard incompatibilist position does not lead to the rejection of morality. The foundational rules of morality proscribe engaging in wrongful conduct but have little to say about how and when judgments of blame are warranted. The hard incompatibilist position is in tension with the practice of blaming others for their conduct, but is compatible with the practice of proscribing wrongful conduct. Therefore, hard incompatibilism can be embraced without jettisoning morality.

---

<sup>326</sup> See, e.g., FLETCHER, *supra* note 22, at 30 (defining wrongdoing as “the incompatibility of the act with a norm . . .”).

<sup>327</sup> Chiesa, *supra* note 97, at 747–48.

<sup>328</sup> See Luis E. Chiesa, *Normative Gaps in the Criminal Law: A Reasons Theory of Wrongdoing*, 10 NEW CRIM. L. REV. 102, 113 (2007); George Fletcher, *What is Punishment Imposed For?*, 5 J. CONTEMP. LEGAL ISSUES 101, 108–09 (1994) (discussing distinction between wrongdoing and culpability).

<sup>329</sup> E.g., George P. Fletcher & Luis E. Chiesa, *Self-Defense and the Psychotic Aggressor*, in CRIMINAL LAW CONVERSATIONS 365, 371–72 (Paul H. Robinson et al. eds., 2009).

<sup>330</sup> See, e.g., *Ellis v. Fixico*, 50 P.2d 162, 164 (Okla. 1935).

<sup>331</sup> E.g., Kenneth W. Simmons, *The Crime/Tort Distinction: Legal Doctrine and Normative Perspectives*, 17 WIDENER L.J. 719, 722 (2008) (noting differences between criminal law and tort law including differences with regard to culpability and treatment of insane persons and children).

A related objection must also be confronted. Can we identify coherent rules of morality if we assume that we lack free will? It seems that we can, at least if we adopt a consequentialist approach to morality.<sup>332</sup> According to consequentialism, the morally correct course of conduct is that which maximizes good consequences.<sup>333</sup> We adopt consequentialist rules in order to influence the behavior of the addressees of the legal rules.<sup>334</sup> Given that humans are capable of understanding moral rules, their conduct is likely to be shaped in some way by the existence of such rules. It is important to note for our purposes that humans possess the capacity to learn and understand rules even if causal determinism obtains and it is deemed to be incompatible with the sort of freedom that undergirds judgments about moral responsibility. As a result, while the hard incompatibilist must concede that her position undermines judgments about blame and about attributing moral responsibility, she is in no way committed to accepting that her position is incompatible with the existence or potential usefulness of moral rules.

##### 5. *Why Assuming that We Lack Free Will May Be Good for Us*

These brief reflections reveal that assuming we lack free will does not drain life of all its beauties and does not deprive us of having healthy and fulfilling interpersonal relationships. While this assumption leads to modifying some aspects of our lives, there is no reason to believe that these modifications would lead to an utterly unappealing conception of life. It might even be the case that the opposite is true. Perhaps our life is made *more* appealing by assuming that humans are not endowed with the sort of freedom that makes them genuinely responsible for their acts. For one, adopting this attitude would inevitably lead to more compassion and understanding towards those who commit transgressions against us. If I truly believe that the person who stole my car is not to blame for his crime, then my reactions to his transgression will likely change from an initial feeling of indignation and resentment, to feelings of understanding and compassion; after all, he is no more to blame for his act than small children are to blame for their conduct or nature is to blame for natural disasters. Assuming that humans lack free will could thus lead to transforming most feelings of resentment and indignation into feelings of compassion and understanding. Living in a world without free will may very well turn out to be quite an appealing prospect.

---

<sup>332</sup> The paradigmatic consequentialist approach to morality is utilitarianism.

<sup>333</sup> *E.g.*, Philip Pettit, *Satisficing Consequentialism*, in 58 PROC. ARISTOTELIAN SOC'Y, SUPPLEMENTARY VOLUMES 139, 165–76 (1984).

<sup>334</sup> *E.g.*, PAUL H. ROBINSON, DISTRIBUTIVE PRINCIPLES IN CRIMINAL LAW 75 (2008); William Shaw, *The Consequentialist Perspective*, in ETHICAL THEORY: AN ANTHOLOGY 463, 470 (Russ Shafer Landau ed., 2007).

## VII. PUNISHING WITHOUT FREE WILL

Even if it is possible to live a fulfilling life without free will, could we still have a functioning and desirable system of criminal justice without free will? Once again, the conventionally accepted answer is that a system of criminal law without free will is likely to be unappealing. The starting point for this argument is that abandoning free will entails abandoning blame. Given that placing blame is an important feature of our current practices of punishment, it may appear that eliminating free will would deal a fatal blow to criminal law. These concerns are overblown. While dispensing with free will means that retributive theories of criminal justice must be abandoned, this does not threaten the consequentialist approaches to punishment. Furthermore, certain consequentialist approaches to punishment that are compatible with the assumption that we lack free will would seem to generate a more humane and efficient system of criminal law than our current system. The purpose of this Part is to explain what punishing without free will would look like and why it is arguably more desirable than a free-will-centered approach to punishment.

A. *Criminal Law Without Retribution: Towards an Incompatibilist and Consequentialist Approach to Punishment*

A criminal law without free will would obviously threaten the role of retribution in justifying punishment. Retribution is the belief that desert is a sufficient condition for punishment.<sup>335</sup> Therefore, retributivists need to ascertain whether someone *deserves* to be punished before they can conclude whether punishing that person is justified.<sup>336</sup> When does a person deserve to suffer under a retributive framework? A person deserves to suffer for doing “X” if, and only if, it is fair to *blame* him for having done “X.”<sup>337</sup> If we assume that humans do not have the sort of freedom that undergirds moral responsibility, it necessarily follows that humans cannot be genuinely blamed for doing what they do. Given that the absence of blame negates genuine desert, retributive practices would thus be unjustified in a world without free will.<sup>338</sup> From a practical standpoint, this means that the mere fact that a rational actor did something evil, bad, or wrongful is not in and of itself a sufficient reason to punish her.

This does not mean that punishment would be unjustified in a world without free will. Although retributive punishment is ruled out by the assumption that humans lack free will, the assumption in no way undermines the imposition of punishment justified on consequentialist grounds.<sup>339</sup> Punishment is justified on

---

<sup>335</sup> See MICHAEL MOORE, *PLACING BLAME* 88 (1997).

<sup>336</sup> See *id.* at 91.

<sup>337</sup> See *id.* (“We are justified in punishing because and only because offenders deserve it.”).

<sup>338</sup> Pereboom, *supra* note 30, at 479–81.

<sup>339</sup> *Id.*

consequentialist aims when the benefits of imposing it outweigh the costs of doing so.<sup>340</sup> The paradigmatic consequentialist justification for punishment is deterring the future commission of offenses.<sup>341</sup> Deterrence, in turn, can be general if it deters the community at large from committing offenses,<sup>342</sup> or specific if it is intended to deter the particular individual who is being punished from committing future crimes.<sup>343</sup> Rehabilitation is another consequentialist justification for punishment that is closely linked to specific deterrence—the aim of rehabilitation is to modify the convicted person’s conduct in a way that decreases the chances that she will recidivate.<sup>344</sup> Finally, incapacitation is a consequentialist aim of punishment usually invoked as a way of neutralizing dangerous offenders who are not candidates for rehabilitation.<sup>345</sup> All of these aims of punishment are not threatened by assuming that actors lack free will, since: (1) the threat of sanction can certainly *influence* the conduct of many actors (i.e., general and specific deterrence) regardless of whether they freely *choose* to be influenced by such forces, (2) behavior modification (i.e., rehabilitation) is perfectly compatible with the assumption that humans cannot freely determine their acts, and (3) the neutralization of dangerous individuals (incapacitation) can be achieved whether the individuals are free actors or not.

Once we shift from retributive to consequentialist punishment, this begs the question: what should trigger the imposition of punishment? Under a retributive account of punishment, a finding of blameworthiness would trigger the imposition of punishment, but such a finding could not justify punishment on consequentialist grounds. Once blame is taken out of the equation, it is unclear what, if anything, could take its place.

One alternative is to shift from blameworthiness to dangerousness as a prerequisite for the imposition of punishment.<sup>346</sup> According to this view, punishment would be justified for the sake of social protection from dangerous individuals. The actor who commits a crime reveals himself as someone who poses danger to others. Punishment would thus be conceived as the state’s way of neutralizing the danger signified by the commission of the offense. In a world without free will, punishment could be reconceptualized as a vehicle for defusing dangerousness rather than as a mechanism for exacting retribution.

---

<sup>340</sup> *E.g.*, R.A. DUFF, PUNISHMENT, COMMUNICATION AND COMMUNITY 3–4 (2001).

<sup>341</sup> ROBINSON, *supra* note 334, at 75.

<sup>342</sup> *Id.* at 8.

<sup>343</sup> *See id.*

<sup>344</sup> *Id.* at 10.

<sup>345</sup> *Id.* at 9–10.

<sup>346</sup> *See id.* at 9.

*B. Punishment as Quarantine (and Other Forms of Societal Protection): More Humane and Efficient than Retributive Based Punishment?*

A good way of illustrating what punishing for dangerousness would look like is to analogize the practice with the way the state fights contagious diseases. When a person is diagnosed with a highly contagious disease, it is sensible for society to take measures to protect itself against the threat.<sup>347</sup> These measures are taken regardless of whether the person who suffers from the disease can be “blamed” for getting sick.<sup>348</sup> The reason for taking these measures is that the sick individual poses a danger to society, rather than blameworthiness for his condition. Similarly, the person who commits a crime under the “punishment as a vehicle to defuse dangerousness” model will often be subjected to certain measures—not because he deserves to suffer, but instead because his conduct reveals that he poses a danger to the community. In the case of sick individuals, the severity of the measures will depend on how dangerous the sickness is believed to be.<sup>349</sup> If the sickness is not particularly dangerous (a common cold, for example), no measures should be required, while if the sickness is sufficiently dangerous it may require that more intrusive measures (vaccination, treatment, quarantine, and so forth).<sup>350</sup> Similarly, while no punishment might be required for a person who has committed a *de minimis* or trivial offense, a person who has committed a serious offense, such as murder and who has done so repeatedly, may require the imposition of much

---

<sup>347</sup> In the United States, the task of protecting society from contagious diseases is of such importance that the government created an agency—the Centers for Disease Control and Prevention (CDC)—to deal with the problem.

<sup>348</sup> While there are certainly cases in which a person is at fault for contracting a disease (think of someone who eats a wild mushroom without first determining whether it is toxic), in many—if not most—cases people contract diseases through no fault of their own (think of someone who takes a plane and is unfortunately assigned a seat next to a person who is very sick and contagious).

<sup>349</sup> *Compare Interim Guidance on Infection Control Measures for 2009 H1N1 Influenza in Healthcare Settings, Including Protection of Healthcare Personnel*, CENTERS FOR DISEASE CONTROL AND PREVENTION (July 15, 2010, 12:30 PM ET), [http://www.cdc.gov/h1n1flu/guidelines\\_infection\\_control.htm](http://www.cdc.gov/h1n1flu/guidelines_infection_control.htm) (“Isolation [of patients infected with H1N1 influenza] should be continued for the [seven] days after illness onset or until [twenty-four] hours after the resolution of fever and respiratory symptoms, whichever is longer . . .”), with *Seasonal Influenza (Flu): Infection Control in Health Care Facilities*, CENTERS FOR DISEASE CONTROL AND PREVENTION, <http://www.cdc.gov/flu/professionals/infectioncontrol/longtermcare.htm> (last updated Sept. 16, 2010) (providing no specific isolation measures, while giving general prevention and hygiene advice).

<sup>350</sup> *Legal Authorities for Isolation and Quarantine*, CENTERS FOR DISEASE CONTROL AND PREVENTION, <http://www.cdc.gov/quarantine/pdfs/legal-authorities-isolation-quarantine.pdf> (last updated Jan. 29, 2010) (authorizing isolation and quarantine measures to prevent the spread of cholera, diphtheria, infectious tuberculosis, plague, smallpox, yellow fever, viral hemorrhagic fevers, severe acute respiratory syndrome (SARS) and any strand of flu that can cause a pandemic).

harsher punishment, including incapacitative measures that would be the criminal law equivalent of quarantine.

One surprising and salutary implication of reconceptualizing punishment as a vehicle for curbing dangerousness is that doing so will likely lead to less reliance on incarceration. Despite the fact that prisons were originally conceived as a place where convicts would go to get rehabilitated,<sup>351</sup> it is widely agreed that prisons are not conducive to rehabilitation.<sup>352</sup> In part because prisons are generally viewed as places that maximize the convict's suffering instead of places where they can get the treatment they need.<sup>353</sup> Thus, imprisonment is the punishment of choice today not because it is effective, but because it allows the state and society to make the convict suffer in order to exact retribution for the harm caused. While this conception of prisons as houses of suffering may adhere to retributive justice theory, it does not adhere to the consequentialist approach to punishment. Under a consequentialist approach, the sanction imposed ought to be the least intrusive sanction that can achieve the desired end in the most economically efficient fashion. Mass incarceration does not fit this standard, since imprisonment is among the costliest and most inefficient forms of punishment.<sup>354</sup> A consequentialist approach to punishment would likely rely as little as possible on incarceration, given that most cases, including drug and weapon offenses, can be dealt with more efficiently through less intrusive sanctions. Examples may include supervised release, probation, drug treatment or intermediate sanctions that are less intrusive than incarceration but more intrusive than standard probation.<sup>355</sup>

There is also good reason to believe that the consequentialist approach to punishment discussed here would push states to assume duties of treatment and rehabilitation of convicts. The analogy between this model of punishment and the state's role in fighting contagious diseases is once again useful to illustrate this point. When someone is diagnosed with a contagious and dangerous disease, the government has three important and distinct obligations. First, the government

---

<sup>351</sup> CURTIS R. BLAKELEY, *AMERICA'S PRISONS: THE MOVEMENT TOWARDS PROFIT AND PRIVATIZATION* 10–11 (2005).

<sup>352</sup> *Id.* at 11.

<sup>353</sup> *See id.* at 12 (“As profit is incorporated into the organizational structure of the prison, it emerges as the dominant goal and becomes the sole measure of success. A cost-to-benefit-analysis is conducted prior to every action, ensuring that the prison will only behave in a manner that proves financially rewarding.”). Compare Adam J. Kolber, *The Subjective Experience of Punishment*, 109 COLUM. L. REV. 182, 200 (2009) (“Experiential-suffering retributivists hold a straightforward view that offenders should be made to suffer in experiential ways; offenders should feel physical and emotional pain and distress.”), with David Gray, *Punishment as Suffering*, 63 VAND. L. REV. 1619, 1650 (2010) (“The error is in the collapsing of two distinct moral concepts—‘crime’ and ‘punishment’—into an undifferentiated category of contingent effects—‘suffering.’ The consequences are far from trivial.”).

<sup>354</sup> STEVEN E. BARKAN & GEORGE J. BRYJAK, *FUNDAMENTALS OF CRIMINAL JUSTICE: A SOCIOLOGICAL VIEW* 456 (2d ed. 2011).

<sup>355</sup> NORVAL MORRIS & MICHAEL TONRY, *BETWEEN PRISON AND PROBATION: INTERMEDIATE PUNISHMENTS IN A RATIONAL SENTENCING SYSTEM* 84–90 (1990).

may defuse the danger presented by the person with the contagious disease.<sup>356</sup> As previously discussed, this can be done in different ways, from prescribing medication to quarantining individuals.<sup>357</sup> Second, the state has a duty to treat the individual in the most effective way available so that he can resume his normal life as soon as practicable.<sup>358</sup> Finally, given that the individual who contracts a contagious disease is not considered evil or blameworthy, the state has an obligation to make the individual's life—while undergoing treatment or quarantine—as pain free and pleasant as possible. Similarly, the “punishment as defusing dangerous offenders” model would lead to three governmental duties, namely: (1) the duty to defuse the danger posed by the wrongdoer by imposing some kind of punishment, (2) the obligation to give treatment to the wrongdoer so that his chances of reincorporating himself to society are increased, and (3) the duty to make the convict's serving of his sentence as pain free and pleasant as possible.

It therefore seems plausible that a consequentialist approach to punishment like, the one advocated here, will lead to a more humane criminal law—shifting the focus of punishment from making people suffer to treating people in a way that maximizes the likelihood that they will once again be able to participate in societal life without posing serious risks to others. With its emphasis on imprisonment, the current system leads to abject prison conditions<sup>359</sup> and an overreliance on incarceration. In contrast, with its emphasis on curbing dangerousness and maximizing good consequences, the consequentialist approach to punishment discussed here would likely lead to less use of prison sanctions and more use of alternative methods of punishment—for example, home detention, treatment for drug and sex offenders, fines, supervised release, and community work.

Increased reliance on these alternative punishments would make our criminal law more humane, compassionate and in tune with our current understanding of human behavior and techniques of behavior modification. Furthermore, this alternative model will likely be more economically efficient than one that relies primarily on incarceration as the preferred type of punishment. Given that imprisonment is a very costly type of punishment,<sup>360</sup> incarcerating offenders should be a measure of last resort. As the literature on the costs of punishment illustrates, we have good reasons to impose imprisonment sanctions only when the risk of detection is extremely low and the offense committed is particularly

---

<sup>356</sup> 42 U.S.C. § 264 (2006) (authorizes the use of measures against contagious individuals in order to prevent the introduction, spread or transmission of communicable diseases).

<sup>357</sup> See *Legal Authorities for Isolation and Quarantine*, *supra* note 350, at 1–2.

<sup>358</sup> Quarantine is typically as a measure of last resort. It can only be used when all other measures fail and should stop being used when least restrictive measures are enough to neutralize the threat. See, e.g., *Frequently Asked Questions About Isolation and Quarantine*, MASS. DEP'T OF PUB. HEALTH 1, 5, [http://www.mass.gov/Eeohhs2/docs/dph/cdc/reporting/iq\\_faq.pdf](http://www.mass.gov/Eeohhs2/docs/dph/cdc/reporting/iq_faq.pdf) (last updated Dec. 2010).

<sup>359</sup> See BLAKELY, *supra* note 351, at 12–13.

<sup>360</sup> BARKAN & BRYJAK, *supra* note 354, at 456.

grave.<sup>361</sup> While such a sparse and economically efficient use of incarceration would be difficult to justify under a retributive criminal law, it would be very easy to justify under the consequentialist model of punishment discussed here.

*C. Objections to Viewing Punishment as a Vehicle for Defusing Dangerousness and Replies to the Objections*

*1. Viewing Punishment as a Vehicle for Defusing Dangerousness Would Justify Punishing the Innocent*

Up to this point, this Article has demonstrated that adopting a consequentialist approach to punishment that eschews any reference to desert and blame as a justification for punishment has many appealing implications. It must be acknowledged, however, that several important objections can be leveled against this conception of punishment. The first and most obvious objection is that any approach to punishment that dispenses with blame could justify punishing the innocent. This is a classic objection to consequentialist approaches to punishment.<sup>362</sup> The retributive model of punishment easily avoids this objection, since the requirement of blame is built in to retributive justice. More specifically, blaming an innocent individual is unjust under a retributive model, for retributive punishment is only justified if the offender deserves to suffer and desert is a function of blame.<sup>363</sup> Consequentialist approaches to punishment cannot avoid the problem easily, for blame is not a prerequisite for punishment under these theories.<sup>364</sup> This is an objection that must be taken seriously, as there is something intuitively troubling about a theory that might justify punishing the innocent.

If punishing the innocent means punishing someone who cannot be genuinely blamed for his conduct, then the consequentialist approach to punishment defended here inevitably leads to punishing the innocent. Given that the point of departure of the theory is that *no one* can be genuinely blamed for his acts, there is no way to escape this conclusion. Nevertheless, this concession is not as damning as it may seem at first glance, for we currently deprive children and the mentally ill of their freedom when they engage in wrongful yet blameless acts. Similarly, the model of punishment defended here would lead to imposing sanctions on blameless individuals who have engaged in wrongful acts. Thus, the only difference between the model defended here and our current approach to punishment is that under the model defended here sanctions imposed on blameless individuals count as “punishment,” whereas under the current model these sanctions are regarded as

---

<sup>361</sup> STEVEN SHAVELL, FOUNDATIONS OF ECONOMIC ANALYSIS OF LAW 551–52 (2004).

<sup>362</sup> See C.L. TEN, CRIME, GUILT AND PUNISHMENT: A PHILOSOPHICAL INTRODUCTION 13–14 (1987) (discussing the “punishing the innocent” objection).

<sup>363</sup> F. Rosen, *Utilitarianism and the Punishment of the Innocent: The Origins of a False Doctrine*, 9 UTILITAS 23, 28 (1997).

<sup>364</sup> *Id.* at 23.



preventative measures. The differences between my proposed model and the current approach to punishment are mostly terminological and not of much normative significance. Furthermore, if punishing the innocent means punishing someone who has not engaged in wrongdoing, then the theory defended here does not lead to punishing the innocent.<sup>365</sup> People who have not engaged in wrongful conduct should not be punished under the “punishment as a vehicle for defusing dangerousness” view, because the occurrence of a wrongful act is the event that triggers the prediction of future dangerousness.

## 2. *Viewing Punishment as a Vehicle for Defusing Dangerousness Would Lead to Punishing People Before They Act*

Another objection to the view of punishment discussed here would be that focusing on dangerousness might lead to punishing dangerous people before they even engage in an act.<sup>366</sup> This objection is inspired in part by Orwellian views<sup>367</sup> of a future that would look like the one depicted in the film *Minority Report*,<sup>368</sup> in which three “pre-cognizant” human oracles can identify future criminals before they engage in the criminal act.<sup>369</sup> This would, in turn, cast doubt upon the so-called act requirement—one of the most venerable and foundational doctrines of criminal law. It might also lead to punishing people for merely contemplating or thinking about committing an offense if such thoughts are believed to be indicative of dangerousness. This is not only normatively unappealing, but also constitutionally suspect. The U.S. Supreme Court has held that imposing punishment on people for so-called status offenses violates the Eighth Amendment proscription of cruel and unusual punishments.<sup>370</sup> Thus, an individual must be punished for what he has done, not for who he is. Furthermore, it is widely

---

<sup>365</sup> There are at least two senses in which an actor might object to punishment based on her innocence. First, the actor might claim that although she engaged in wrongdoing she is innocent because she cannot be fairly blamed for engaging in the admittedly wrongful act—that is that her wrongdoing should be excused. Second, the actor might claim that she is innocent because she did not engage in a wrongful act at all. The theory of punishment espoused in this section can be objected based on the former conception of “innocence,” but not the latter.

<sup>366</sup> Robert Batey, *Minority Report and the Law of Attempt*, 1 OHIO ST. J. CRIM. L. 689 (2004).

<sup>367</sup> GEORGE ORWELL, *NINETEEN EIGHTY-FOUR* 174 (Holt Rinehart & Winston 2000) (1949) (“Thoughts and actions which, when detected, mean certain death are not formally forbidden, and the endless purges, arrests, tortures, imprisonments, and vaporizations are not inflicted as punishment for crimes which have actually been committed, but are merely the wiping-out of persons who might perhaps commit a crime at some time in the future.”).

<sup>368</sup> *MINORITY REPORT* (20th Century Fox Film Corporation, DreamWorks SKG 2002) (Based on a short story written in 1956: PHILIP K. DICK, *The Minority Report*, in *THE MINORITY REPORT AND OTHER CLASSIC STORIES* 71 (2002)).

<sup>369</sup> *Id.*

<sup>370</sup> *See, e.g., Robinson v. California*, 370 U.S. 660, 666–67 (1962) (“Even one day in prison would be cruel and unusual punishment for the ‘crime’ of having a common cold.”).

believed that the Constitution also proscribes punishing mere thoughts or internal deliberations in the absence of some sort of act that demonstrates the firmness of their intentions.<sup>371</sup>

The first reply to this objection is that there is currently no machine or human that can predict with accuracy whether a given individual will commit a crime in the future.<sup>372</sup> Therefore, the risk of punishing someone who is not really dangerous based on abstract predictions of dangerousness is quite high. Given that the harm caused as a result of punishment in such cases is quite significant (deprivation of freedom in many cases), we have good reasons to inflict such harm only if there is near certainty about the triggering conditions that justify its infliction (dangerousness, according to the theory of punishment discussed here). Since we currently do not have the capacity to predict with any degree of certainty whether particular individuals who have not committed a wrongful act in the past will behave dangerously in the future, we ought to abstain from punishing people before they act based on predictions of dangerousness that are not grounded on the commission of prior wrongful acts. But what if sometime in the future we develop *Minority Report*-type technologies that allow us to predict future criminal acts with great accuracy well before that person commits an offense?<sup>373</sup> Would it be acceptable to punish those persons before they engage in the wrongful conduct? If these cases ever arise in the future (this is a very big “if”), it might be legitimate to take action against the individual before he acts, and in some extreme cases limiting the actor’s freedom might be warranted. However, it might be better to treat these cases as instances of civil confinement, for it is not clear whether it makes sense to talk about “punishing” someone who has yet to do anything wrongful.<sup>374</sup>

#### D. Punishing Without Free Will—Summary

The conventional wisdom is that it would be normatively unappealing to assume that humans lack free will. Making such an assumption would lead to an impoverished view of life and an unattractive system of criminal justice. Despite

---

<sup>371</sup> See, e.g., *Stanley v. Georgia*, 394 U.S. 557, 566 (1969) (“Our whole constitutional heritage rebels at the thought of giving government the power to control men’s minds.”).

<sup>372</sup> Some jurisdictions currently use a computer program that can help predict whether certain convicts are more or less likely to recidivate. Professor Richard Berk of the University of Pennsylvania developed the program. Professor Berk acknowledges, however, that the program “[is not] anywhere near being able to do [what the *Minority Report* pre-cognizants could do].” Daniel Bates, *The Real Minority Report: U.S. Police Trial Computer Software That Predicts Who Is Most Likely to Commit a Crime*, DAILY MAIL ONLINE, <http://www.dailymail.co.uk/sciencetech/article-1306070> (last updated Aug. 25, 2010 5:44 PM).

<sup>373</sup> Media reports are already drawing parallels between Professor Berk’s program and the machines used in *Minority Report* to predict future crimes. *Id.*

<sup>374</sup> This, in a sense, is something we already do when we decide we deal with a case of civil confinement.

its intuitive appeal, there are good reasons to believe that the conventional wisdom is wrong. Assuming that humans lack free will would lead to eliminating retribution as a justification for the imposition of punishment. Contrary to what avowed retributivists would have us believe, discarding retribution does not make our criminal justice system less attractive. A criminal law that does not rely on retribution as a justification for punishment ought to conceive punishment as a way of neutralizing dangerous offenders. Conceptualizing criminal law this way is likely to lead to a more economically efficient and humane system of criminal justice that relies less on incarceration and more on treatment and rehabilitation. As a result, there are good reasons to believe that assuming that humans lack free will would generate a more normatively appealing criminal law than the one we have today and this, in turn, provides us with good reasons to embrace an incompatibilist solution to the free will problem.

#### VIII. CONCLUSION

This Article opened with a reference to Voltaire's *Candide*.<sup>375</sup> In this celebrated book, Voltaire called our attention to the problem of free will by contrasting Candide's view that what makes humans special is a kind of freedom that is lacking in other animals, with Martin's view that determinism rules out the possibility of free will and, perhaps, moral responsibility. Since *Candide* was published many philosophers, scientists, and legal theorists have proposed different ways of solving the free will problem. After all this time, it can be asserted with confidence that there is no "knockout" argument that demonstrates that free will is compatible or incompatible with the truth of causal determinism. In the face of this dialectical stalemate, this Article has argued that we ought to decide whether to hold on to or abandon free will by asking whether assuming that we have free will is more or less normatively appealing than making the opposite assumption. This Article suggests that there are good reasons to believe that the most normatively attractive way out of the free will maze is to assume the incompatibility of determinism and indeterminism with the type of freedom that lies at the core of contemporary criminal law and theory. Doing so would not make our lives any less appealing, and it could lead to a more humane and efficient system of criminal justice than one that assumes that we are endowed with the sort of free will that undergirds judgments of blame and praise. Consequently, contrary to what most philosophers and criminal scholars argue, punishing without free will is not less appealing than the alternative.

---

<sup>375</sup> VOLTAIRE, *supra* note 1; *supra* text accompanying note 1.