

## Essay

### Free Will and Law: Toward a Pragmatic Approach

Greg Simmons

Despite its profound significance for notions of legal responsibility, the courts and legal system have tended to avoid direct engagement with the philosophical problem of free will. Focusing on mental illness and the criminal law, I advance here a naturalistic approach that builds on the work of P.F. Strawson, one I believe offers a pragmatic basis from which to address the contradictions and challenges present when folk wisdom, science, philosophy and the law intersect. In this way, I contend that moving dialectically between a reflexive engagement with extant practical attitudes to freedom and the empirical investigation of the participant/object divide affords the opportunity to develop more rational and humane legal and social responses to both the mentally disordered and broader population.

#### Introduction

In everyday life, most of us tend to operate under the assumption that we are free to make our own choices. Whereas our range of options may be physically or socially limited, the decision of which to pursue is, in most circumstances, our own. At times, however, we encounter deeper constraints on freedom that seem to place internal restrictions on individuals – addictions, phobias, or mental illness, for example. Thus, we also tend to assume that at some point physiological or mental conditions can override our normal capacity for free decision-making. A similar perspective is often present in the clinical setting, from where it spills into the legal treatment of mental illness and intersects with the law's own assumptions about what it means to be a legally responsible agent. At the philosophical level, however, the problem of free will appears as a (thorny) issue of general application, one most often linked to determinism.

I first examine here the implications of this philosophical problematic for the legal treatment of mental illness and criminal responsibility, before putting forward a naturalistic approach that is based on the work of P.F. Strawson – one I believe offers a pragmatic basis from which to address the contradictions and challenges present when folk wisdom, science, philosophy and the law intersect. Rather than undertake a detailed jurisprudential analysis, my engagement with legal doctrine is at a general level, using Canadian law as a reference, to the end of addressing how philosophical engagement with the problem of free will should inform legal and criminal justice policy. I compare and contrast my analysis with the work of prominent legal theoreticians, Stephen J. Morse and Michael S. Moore. Whereas these scholars each make a case that at the doctrinal level the criminal law can operate independently of the philosophical problematic of free will, I contend that empirically the issue is threaded through legal discourse (if much of the time only implicitly). I also aver that, more importantly, a pragmatic resolution of

the problem is instructive in forging a legal and criminal justice system that is truly just. I also note at the outset that, in looking to the legal implications of philosophical inquiry, I do not follow the standard route of philosophical investigation: rather than unidirectionally deriving from the free will issue implications for responsibility and choice, I seek to move dialectically between attitudes to freedom as they currently exist and philosophical arguments as to what they should be. An iterative process is already to some degree evident in the ambivalent approach of the legal system to freedom and responsibility, and I will attempt to show how making it an explicit methodological tack affords the opportunity to develop a more rational and humane treatment by the legal and criminal justice system of both the mentally disordered and the broader population.

### **Determinism and the Problem of Free Will**

Determinism is most often formulated in causal terms: roughly stated, every event or state of affairs is causally necessitated by preceding events or states of affairs. Put nomologically, *natural laws* determine the course of events in the world.<sup>1</sup> Prior to the advent of modern physics, a deterministic worldview underwrote scientific inquiry in general, perhaps reaching its strongest form in the nineteenth century positivism. The development of quantum mechanisms in the early part of the twentieth century, however, shook this bedrock conception of a wholly deterministic universe, revealing a fundamental indeterminism at the level of particle physics.<sup>2</sup> Nonetheless, at the macro-level, the mass resolution of quantum level probabilities leads us to observe regularized spatiotemporal contiguities between events such that in our everyday world the

---

Sincere thanks to Professor Simon Verdun-Jones for comments on an earlier version of this paper.

<sup>1</sup> An example of a more precise definition is that offered by Van Inwagen as the conjunction of these two propositions:

- (a) For every instant of time, there is a proposition that expresses the state of the world at that instant.
- (b) If A and B are any propositions that express the state of the world at some instants, then the conjunction of A with the laws of physics entails B.

See P Van Inwagen, "The Incompatibility of Free Will and Determinism" (1975) 27 *Philosophical Studies* 185 at 186. Determinism can also be framed in epistemological terms in reference to predictability. Simon-Pierre Laplace posited that an entity that knows the exact conditions of the universe at any particular moment and all the natural laws that govern the universe would be able to predict the future, down to the smallest detail (Pierre Simon Laplace, *A Philosophical Essay on Probabilities*, 6th ed, translated by FW Truscott & FL Emory (New York: Dover Publications, 1951) at 4). The nature of the laws of nature and of causality, of course, are long-standing and expansive philosophical problems – problems that will not detain us here.

<sup>2</sup> This is not a wholly settled matter. There are deterministic interpretations of quantum mechanics. See, for example, Hrvoje Nikolic, "Bohemian Particle Trajectories in Relativistic Bosonic Quantum Field Theory" (2004) 17 *Foundations of Physics Letters* 363.

universe at least appears to operate deterministically. And beyond the domain of particle physics, determinism has retained an axiomatic role in the sciences, both soft and hard.

Once the causal necessity of physical law is applied to mental states, determinism raises sharp concerns for the freedom of human action in general, and responsibility and morality in particular. If I am wholly instantiated in a deterministic world, then everything that goes on “in my head” – my thoughts, desires, beliefs and attitudes – is the product of a causal chain extending eventually back to forces and situations beyond my control (to some interplay of genetic and environmental factors). How, then, are such mental states really my own? And if, in the final analysis, the decisions I make on the basis of them (and that are themselves such states) are merely the result of events unfolding according to the laws of nature, how have I made a genuine choice at all? I no more choose to do what I do than the ball “chooses” to fall to the floor under the influence of the law of gravity. Natural laws being natural laws, I cannot act in violation of them – meaning, I could not have acted other than I did. Further, precluding the counterfactual possibility that lies at heart of a free choice means that the determinist threat extends to the realm of ethics: if I cannot choose otherwise, how can I be held morally responsible for my actions? But if determinism is not true, if our decisions are not caused, then how are they any different from random, meaningless events? How can I be said to make morally responsible choices, or, in fact, choices that are meaningful at all?

Various positions can be adopted in response to the challenge of determinism. The view that we indeed live in a deterministic universe and that this is fatal to our intuition that we have free will is termed *incompatibilism* (or *hard determinism*). Incompatibilism sees both counterfactual possibility as necessary to free will and antecedent causal sufficiency as blocking this very possibility. As a result, the sense that we could act other than we do is merely an illusion, itself the result of causal antecedents that create the impression that we are immersed in a world of possibility but lead in fact to a particular necessary outcome.<sup>3</sup> Some incompatibilists contend that deterministic biological systems are not ontologically distinct from deterministic mechanical systems; and because these systems do not possess free will, neither do we.<sup>4</sup> Another argument is that true freedom requires a breaking of the causal chain of events at the level of the

---

<sup>3</sup> See, for example, Derk Pereboom, *Living without Free Will* (Cambridge: Cambridge University Press, 2001).

<sup>4</sup> See, for example, John M Fischer, “Incompatibilism” (1980) 43 *Philosophical Studies* 127. Whether alien systems can possess free will or various other properties of mind is a matter of active debate. One of the criticisms of type-physicalist theories of mind – theories that ascribe a one-to-one correspondence between mental states and physical states – is that they necessitate an essentialist equating of our mental states with our specific physiological make-up. The result is that creatures of similar structural complexity to ourselves, but who are silicon-based rather than carbon-based, for example, could not possess intentionality, consciousness, etc. Other physicalist monist theories of mind – token identity or functionalist theories that see mental states as residing in the organizational structure (or “software”) of the mind – recognize the potential for non-human creatures or machines to have mental states very similar to our own. See, for example: Gary Young, *Philosophical Psychopathology: Philosophy without Thought Experiments* (Hampshire, UK: Palgrave Macmillan, 2013).

agent: she must be the “originating” or “ultimate” cause of her actions. Since determinism precludes this, free will is impossible.<sup>5</sup> The consequences of the reality of determinism are that the moral and normative frameworks that underpin society are a foundation of sand and a radical reworking of our moral and legal systems is required in order to place them on a coherent basis.

*Compatibilism*, in contrast, sees determinism and free will as reconcilable. Usually proceeding from a physicalist monist philosophy of mind,<sup>6</sup> compatibilists imply that to hold that causation negates freedom is to commit what philosophers call a category error. Freedom does not require some kind of metaphysical power to act outside of the laws of nature; we are sufficiently free if we are not constrained or compelled – that is if we can voluntarily act upon our desires and preferences.<sup>7</sup> Barring such restrictions, we are free to choose other than we do, despite that this would require different causal antecedents. Rather, it is the existence of these specific antecedents that makes our choices meaningful. It is through them that our action becomes coherent and explainable as opposed to merely random. Responsibility, in turn, amounts to the capability to respond to rewards and punishments and the incentives of praise and blame.

*Libertarianism* accepts the incompatibility of determinism and free will, but holds that we have free will and so determinism is false.<sup>8</sup> The fact that human beings are able to act other than they do reflects our indeterministic and self-reflexive nature, a definitive characteristic of what it means to be human.<sup>9</sup> Responsibility requires more than the compatibilist’s claim of responsiveness to incentives and disincentives present in the environment. For libertarians, the compatibilist’s utilitarian approach misses the deeper sense of responsibility that comes from

---

<sup>5</sup> See, for example, Robert Kane, *The Significance of Free Will* (Oxford: Oxford University Press, 1998).

<sup>6</sup> While the reductionism that usually accompanies physicalist monism may throw the deterministic challenge in particularly stark relief, it is not necessary to determinism per se. If psychological or sociological laws have causal effect, then they could operate deterministically even if they are not fully reducible to physical laws.

Further, it is conceivable to have an idealist or dualist conception of determinism in which mental states as non-physical substances operate through causal sufficiency to produce particular subsequent mental states to the preclusion of other potential ones. This would offer a solution to some of the problems for physicalist theories of mind – the problem of *qualia* and the distinctly subjective, phenomenological quality of what it is *like* to have a specific experience, for example, something that seems hard to reduce to physical states. Traditional problems for dualism and idealism, such as ascertaining the nature of this non-physical substance, would remain – as would the problem of free will (the question now being how one could have acted contrary to dictates of prior mental – rather than physical – states).

<sup>7</sup> Classic early exponents of compatibilism in modern western philosophy include Thomas Hobbes and David Hume.

<sup>8</sup> In this regard, libertarianism is actually a form of incompatibilism, one that reconciles the incompatibility between free will and determinism by denying the latter rather than the former.

<sup>9</sup> See, for example, Robert Kane, “Free Will: New Directions for an Ancient Problem” in Robert Kane, ed., *Free Will* (Malden: Blackwell, 2001).

truly making your own decisions in a way not bound by deterministic physical law. Libertarianism can take a dualist form, in which a non-physical entity or substance (perhaps a soul or spirit) acts outside of or overrides physical causality.<sup>10</sup> Alternatively, the contra-causal power may reside in the material universe, making for a version of libertarianism that is consistent with a physicalist monism. Some libertarians point to the indeterministic understanding of the universe provided by quantum mechanics as opening a space where freedom might reside and, through probabilistic laws, interact with physical reality.<sup>11</sup> Yet, whereas attempts have been made to show specifically how quantum mechanics might convey upon a sentient being a freedom not otherwise available,<sup>12</sup> the connection remains far from clear. The difficulty that libertarians face is to provide an account of free will that extends beyond the negative claim that our actions are not causally necessitated. Capturing this more profound sense of contra-causal freedom in a coherent manner is the preeminent challenge for libertarianism. If it is not through connection to causal antecedents, then under what process do our decisions become meaningful as opposed to being merely random, uncaused events? Perhaps there is an agency at work that, while remaining influenced by causal precursors, can somehow override them; but its nature remains opaque.<sup>13</sup>

### **Some Implications of Neuroscience for Free Will**

Recent advances in neuroscience have heightened the focus on the relationship between the brain and the mind, revealing specific physical correlates or instantiations (depending on the philosophical theory of mind adopted)<sup>14</sup> of mental processes. And while they by no means settle the free will debate, increasingly sophisticated understandings of the structure and functioning of

---

<sup>10</sup> René Descartes was a classic early exponent of this view. See René Descartes, “Passions of the Soul”, *The Philosophical Works of Descartes*, vol 1, translated by ES Haldane & GRT Ross (London: Cambridge University Press, 1931).

<sup>11</sup> See, for example, Robert Nozick, *Philosophical Explanations* (Cambridge: Harvard University Press, 1981) at 298.

<sup>12</sup> See, for example, Sabine Hossenfelder, “The Free Will Function” (2012) ArXiv12020720 Phys.

<sup>13</sup> The fourth logical position, that human beings are indeterministic but do not have free will, is not widely advanced. But if determinism is accepted and the libertarian cannot meet the challenge of reconciling *it* with free will, then the essential randomness of the universe could be seen as foreclosing the possibility for true freedom.

<sup>14</sup> Eliminative materialism sees brain states as completely explaining mental states, with the result that our common sense understanding of the latter is wholly mistaken. Other physicalist theories (token identity theories or functionalist theories) see brain states as sufficient but not necessary instantiations of mental states, such that the same mental states could arise from other types of physical forms. Emergentism is a non-reductive physicalism holding that structures or systems of sufficient complexity may exhibit qualities that cannot be reduced to lower level components. Just as water has properties distinct from its component atoms of hydrogen and oxygen, the brain manifests supervening properties (in the form of mental states) that cannot be wholly located in the sum of neurons, synapses or other component parts.

the brain and their correlation with mental states and processes is often taken as lending increased support to physical determinist views of the mind. Neuroscience undoubtedly offers great value, and transposed to the criminal justice setting its findings can lead to a refining, clarifying and ultimately potentially humanizing of the treatment not only of the mentally ill but of offenders more generally. In regard to the philosophical issue of free will, however, its implications are less straightforward and highly dependent on context.

Many of these research efforts are oriented toward identifying relevant differences in pathological or abnormal individuals or populations. The result can be a tendentialism toward uni-causal and reductionist explanations. The scientific study of human beings inevitably involves a shift in perspective away from the subject-orientation of everyday human interaction and toward an object-orientation inherent to systematic empirical investigation. The result is that deterministic assumptions regarding human cognition and behaviour are featured in a way that they are not in our ordinary explanations of what is “going on in people’s heads.” This is not to say that researchers cease to view participants as human, only that the precepts of scientific investigation require a focus on abstracted properties amenable to measurement, quantification and theoretic reduction. Arguably, when the object of study is those identified as abnormal and the means is the measuring or visual imaging of biological structure or activity, this tendency to reductionist explanations rooted in physical determinism is still further reinforced.<sup>15</sup>

The challenge that the thesis of causal determinism presents, however, is to free will, generally. Individuals with abnormal brain structure or function are no *more* determined than those without: as a metaphysical model of the world, determinism does not submit to degrees. Abnormal brain structure or function may limit (or expand) individual freedom but it does so not by imparting deterministic restraints on an otherwise contra-casual system. It may, however, restrict the range of available choices, generally, and the range of choices that society deems acceptable or appropriate even more so. In other words, although there is no fundamental difference in ontological status between mentally “normal” and “abnormal” individuals, there may be very relevant distinctions of *kind*. In this regard, neurological and biological descriptions – even of a highly reductionist nature – remain meaningful and pertinent. It is only the temptation to link them to basic metaphysical differences between human beings that remains misplaced.

## Free Will and the Law

---

<sup>15</sup> Certainly, many commentators seem keen to map particular aspects of human behaviour to particular anatomical features of the brain in a one-to-one manner. (See, for example, Knabb et al “Neuroscience, Moral Reasoning, and the Law” (2009) 27 Behavioral Sciences and the Law 219.) This tendency risks eliminating explanations that may not actually be fully reducible to more “fundamental” theoretical levels. It may be, for example, that the functionalism present in biological explanations is particularly relevant to explaining the operation of the brain and neurological system, yet it is eliminated in an overly isolationist anatomical approach.

How are issues of free will reflected in the legal realm? One position, adopted by the legal theorist, Stephen Morse, is not at all. While recognizing that the issue of free will is a genuine (and profound) philosophical problem, Morse contends that the concept itself is not embodied in any legal doctrine, nor is it a foundational of criminal responsibility. In determining responsibility, the law certainly considers problems pertaining to “consciousness, the formation of mental states such as intention and knowledge, the capacity for rationality, and compulsion, but it never addresses the presence or absence of free will.”<sup>16</sup> In short, the claim is that the predicates of criminal responsibility – actions and mental states – can be given intelligible, consistent and coherent meaning, independently of the truth or falsity of determinism and its philosophical implications.<sup>17</sup> Like Morse, Michael S. Moore has devoted considerable efforts to distinguishing the issue of free will from that of criminal responsibility: legal excuses are not based in causation, rather in the absence of the ability, capacity or opportunity to do what is legally required.<sup>18</sup> However, while Moore is of a mind with Morse that causation is not a grounding for excuse, he nonetheless contends that the operation of excuse and responsibility must still be understood within the philosophical framework of free will (49). For instance, he notes the hard determinist challenge that if causation necessitates excuse and causation is universal, then everything is excused cannot simply be side-stepped by those seeking to distinguish excuse from causation. The response that the result is an absurdity only holds if we see excuse and responsibility as meaningful and distinct, but this is the very position that is the target of the hard determinist attack. Further, Morse’s argument still depends on the possibility of the counter-factual – so again a capacity-based account of excuse cannot simply bracket the hard determinist challenge that the reality of determinism precludes the possibility of acting otherwise and hence of true choice.<sup>19</sup> Moore claims the appropriate – and only fully coherent – response to the philosophical problem of free will is compatibilism and that it is within such an understanding that a capacity-based explanation must be nested.<sup>20</sup> For both Morse and Moore, neuroscience can inform our understanding of human freedom but unless and until it reaches a level of development such that it undermines the law’s “folk psychological” understanding of human behaviour (that looks to desires, intentions, beliefs, etc.) it will not override legal explanation or principle.<sup>21</sup> Legal excuses are premised in the undermining not of free will but of

---

<sup>16</sup> Stephen J Morse, “Avoiding Irrational NeuroLaw Exuberance: A Plea for Neuromodesty” (2011) 3 *Law, Innovation and Technology* 209 at 215. Michael S Moore, “Responsible Choices, Desert-Based Legal Institutions, and the Challenges of Contemporary Neuroscience” (2012) 29 *Social Philosophy and Policy* 233.

<sup>17</sup> Morse’s own position is compatibilism, which he contends is consistent with the legal principles relevant to criminal responsibility.

<sup>18</sup> For example, see Michael S Moore, “Causation and Excuses” (1985) 73 *Cal L Rev* 1091 at 1128-37.

<sup>19</sup> See Michael S Moore, “Stephen Morse on the Fundamental Psycho-Legal Error” (2016) 10 *Criminal Law and Philosophy* 45.

<sup>20</sup> *Ibid.*

<sup>21</sup> *Supra* note 18. The primary function of the law is as a guide to action. Hence, the law operates on the folk psychological realm of reasons and intentions. Whereas the mechanistic approach of neuroscience, at

the capacity for “practical reason” – be it through duress, necessity, involuntariness, compulsion, addiction or insanity.

Moore and Morse each make a cogent case that, doctrinally, positive law can function independently of the philosophical problem of free will<sup>22</sup> – that, in essence, legal impairments to the capacity for choice can be distinguished from the broader issue of what the truth or falsity of determinism may or may not mean for such a capacity.<sup>23</sup> Nonetheless, empirically, it remains the case that assumptions regarding causation and human freedom are interwoven into the operation of the legal system and the pragmatic reasoning of judges. Factually, the folk psychological

---

least at its present level of development, cannot fulfil this normative role, it may, nonetheless, help in the application of legal doctrine, e.g., in determining whether a genuine excuse existed or potentially inform rehabilitative measures. See also, Stephen J Morse, “Determinism and the Death of Folk Psychology: Two Challenges to Responsibility from Neuroscience” (2008) 9 Minn. JL Sci & Tech 1. Morse is critical of the legal system for being too often overly deferential to scientific expertise. (See, for example, Stephen J Morse, “The New Syndrome Excuse Syndrome” (1995) 14 Criminal Justice Ethics 3. Here he agrees with Robin Feldman, who sees strategies of both “internalization” and “externalization” at work. In the former, the law adopts scientific criteria as legal criteria, while in the latter, legal decision-making is essentially outsourced to experts from other fields. See Robin Feldman, *The Role of Science in Law* (New York: Oxford University Press, 2009).

<sup>22</sup> Both Moore and Morse expound their case through extensive and often subtle argumentation, a full engagement with which is beyond the scope and aims of this paper. See, for example: Stephen J Morse, “Avoiding Irrational NeuroLaw Exuberance: A Plea for Neuromodesty” (2011) 3 Law, Innovation and Technology 209 at 215; “Determinism and the Death of Folk Psychology: Two Challenges to Responsibility from Neuroscience” (2008) 9 Minn. JL Sci & Tech 1; “The Non-Problem of Free Will in Forensic Psychiatry and Psychology,” (2007) 25 Behavioral Sciences and the Law 203; “Uncontrollable Urges and Irrational People Essay” (2002) 88 Va L Rev 1025; “The New Syndrome Excuse Syndrome” (1995) 14 Criminal Justice Ethics 3; “Treating Crazy People Less Specially” (1988) 90 W Va L Rev 353; “Failed Explanations and Criminal Responsibility: Experts and the Unconscious” (1982) 68 Va L Rev 971; Michael S Moore, “Stephen Morse on the Fundamental Psycho-Legal Error” (2016) 10 Criminal Law and Philosophy 45; “Responsible Choices, Desert-Based Legal Institutions, and the Challenges of Contemporary Neuroscience” (2012) 29 Social Philosophy and Policy 233; “Causation and the Excuses” (1985) 73 Cal L Rev at 1091; “Responsibility and the Unconscious” (1980) 53 S Cal L Rev 1563; “Failed Explanations and Criminal Responsibility: Experts and the Unconscious” (1982) 68 Va L Rev 971;

<sup>23</sup> A classic statement of capacity as the basis for responsibility (or the lack thereof) is found in Hart: What is crucial is that those whom we punish should have had, when they acted, the normal capacities, physical and mental, for doing what the law requires and abstaining from what it forbids, and a fair opportunity to exercise these capacities. Where these capacities and opportunities are absent, as they are in different ways in the varied cases of accident, mistake, paralysis, reflex action, coercion, insanity, etc, the moral protest is that it is morally wrong to punish because “he could not have helped it” or “he could not have done otherwise” or “he had no real choice”. HLA Hart, *Punishment and Responsibility* (Oxford: Clarendon Press, 1968) at 152.

concepts upon which the law operates include notions of causal impediments (by insanity, addiction and so on) to free choice. Further, at an ontological level, the question of free will is central to what it means to be a moral agent and responsible actor, understandings of which are also intrinsically bound up in the pragmatic operation of legal liability. A realist understanding of the law recognizes that legal doctrine is informed and shaped by other discourse and systems of knowledge – as, indeed, it should be. Judges and lay-people alike may with Moore and Morse view legal excuse as based in constraints on the *capacity* for choice, yet it is clear that the attenuation of capacity is often attributed to a causally-based undermining of otherwise contra-causal freedom. Hence, Morse and Moore’s own prioritizing of the folk psychological as the descriptive and normative basis of the law’s operation, means (ironically) that a folk psychological understanding of free will cannot be extricated from a realist analysis of legal doctrine.

Below I proffer an account of free will grounded in the work of P.F. Strawson that is both explanatory and normative – with elements pertaining to how the legal system does and should engage with the issue. With Moore and Morse, the issue of capacity plays a central role. Also with Moore and Morse, folk psychological concepts and a pragmatic approach to legal concepts are pertinent, but by way of contrast, an appeal to a semantic reductionism that excludes free will is avoided: factually and counter-factually, the conceptualization and attempted resolution of the problem of free will matters to how the law and the legal system both do and should behave. Hence, it is appropriate to first, briefly, consider the practical role the concept of free will plays within legal discourse.

### *Actus reus and Voluntariness*

To establish the *actus reus* of an offence the Crown must prove that the conduct in question was undertaken voluntarily – if it was involuntary, the *actus reus* is abrogated. Courts have long held that finding someone criminally responsible for acts not voluntarily chosen would be fundamentally unjust. Conduct can be involuntary when external events conspire to place the accused in a situation not of her own making, but the more metaphysical sense of involuntariness pertains to the legal concept of *automatism*, in which an accused does not have conscious control over her actions with the result that she is not making a willed or deliberate decision. In so doing, the law distinguishes various kinds of automatism, with various legal outcomes. Automatism that does not result from a mental disorder, for example, may lead to a complete acquittal, but when a mental disorder is a causal factor, the outcome is a verdict of Not Criminally Responsible on Account of Mental Disorder (NCRMD).<sup>24</sup> Traditionally, automatism as a consequence of self-

---

<sup>24</sup> Stone (1999) 154 CCC (3d) 353. The defence of mental disorder is codified in section 16 of the *Criminal Code*, R.S.C. 1985, c. C-46 which states, in part:

induced intoxication would not fully exonerate the accused but could be raised as a partial defence, resulting in conviction for a less serious included offence.<sup>25</sup> In spite of the variation in attendant legal culpability, courts tend to talk of automatism as premised on the notion of an impaired will – an inability to choose or to intentionally act – as distinguishing the autonomic actor from the conscious, normally functioning individual.<sup>26</sup> It may be that – as Morse and Moore claim – what is actually meant here is the absence of the capacity for rational choice (and the courts are just using “free will” as a short-hand for criminal). Nonetheless, the judicial foregrounding of the causal connection between a substance or state and this resulting negation of choice stands in contrast to the default position in which the ability to choose how to act is taken as given and consideration of causal antecedents not brought to bear. Thus, there is a, generally implicit, suggestion of a contra-causal free will undermined by whatever agent or condition has induced the automatist state – even if the philosophical basis for this remains largely unaddressed.

At other times, however, judicial conceptions of involuntariness seem more consistent with a compatibilist sense of free will, in particular, those that look to the absence of conscious awareness as (sometimes) absolving the accused of responsibility.<sup>27</sup> The degree to which consciousness need be impaired is not fully clear. Psychiatric evidence suggests that individuals in automatistic states are close to unconsciousness, their activity driven by reflex or the playing out of routine behavioural scripts in a manner somewhat akin to a movie zombie. Legal findings of automatism have been made in situations involving more elaborate behavior that, *prima facie*, looks to be purposeful, however. In *Parks*,<sup>28</sup> for example, the accused was found to have been in a somnambulistic state while engaging in seemingly deliberate and purposive action – driving 23 kilometres to the residence of his in-laws and proceeding to beat and stab them.

---

16 (1) No person is criminally responsible for an act committed or an omission made while suffering from a mental disorder that rendered the person incapable of appreciating the nature and quality of the act or omission or of knowing that it was wrong.

<sup>25</sup> The somewhat complicated state of Canadian law at present is that extreme self-induced intoxication that induces a state “akin to automatism or insanity” could be a complete defence to charges that do not involve personal violence or the threat thereof. For these offences, self-induced automatism at most amounts to a partial defence. (See *Daviault* [1994] 3 SCR 63 and the Parliamentary response of section 33.1 of the *Criminal Code*.)

<sup>26</sup> In limiting automatism as a defence in the case of self-induced intoxication, the courts are looking to a broader situational culpability that invokes responsibility for the conditions leading to the inducement of the automatistic state. Free will retains its relevance in the accused’s choice to willingly put himself in a state in which criminal behaviour was a foreseeable possibility.

<sup>27</sup> It should be noted that a lack of consciousness is not a necessary condition for involuntariness more generally. In the case of a reflex response, for example, the individual may be completely conscious and aware of what is happening, the key element being a lack of control.

<sup>28</sup> *Parks* [1992] 2 SCR 871, 75 CCC (3d) 387.

In fact, the nature of consciousness and what it means to speak of “impaired” consciousness is a vast topic in itself, and a thorough consideration is well beyond the scope of this exercise. It is worth lending some consideration here to the role of consciousness in relation to the free will debate, however. Upon first blush, it may seem reasonable to hold that if we are not conscious of our actions that they cannot be free. Yet we routinely engage in goal-directed, meaningful behaviour that is not the result of a higher level reflective process but which seems no less free for its absence – driving a well known route home while thinking about dinner, for example. Likely, the “freedom” of this action rarely comes into question because it is routinized behaviour subordinate to an intentional and deliberate goal made at a more fully conscious level. In contrast, the claim that a sequence of events that under normal circumstances would suggest intentional action in furtherance of a deliberate plan was actually executed in an automatistic states (as in *Parks*) strikes us as at odds with our normal understanding of human behaviour and cognition. If determinism holds, then both types of action are equally determined, yet issues of freedom and responsibility seem to arise in one case but not the other, something that hints at tacit normative and situational factors relevant to our understanding of “willed” action. In the latter case, freedom has been assailed by the breakdown of the normal holistic functioning of the “self,” the ability to draw on memories and process and integrate them with desires, beliefs and emotional content to form various levels of intention and intentionality. It is this very kind of destabilization of the self as a cohesive system that characterizes the “dissociative state” indicative of automatism (at least of some kinds) and during which the individual is – literally – “not herself.” Here capacity does seem to be the critical factor. It would appear that it is this varying capacity for choice – open, yet to differing degrees neurologically, psychologically and sociologically circumscribed and all the while located contextually and normatively within a social environment – that meaningfully defines the nature and extent of freedom.

### *Mens rea and intentional choice*

Similar ambiguities are manifest in the *mens rea* element of an offence. While the exact mental elements constituting *mens rea* vary with the offence, in all but strict liability offences (which do not attract the moral stigma of “true crimes”), the ability to form intent is an underlying presupposition.<sup>29</sup> The result is that the classical problem of free will bears upon the criminal law’s concept of the “guilty mind.” A subjective *mens rea* standard requires proof that the accused either deliberately intended to bring about the proscribed consequences of her action or subjectively knew that they likely would result. Objective *mens rea*, on the other hand, is concerned only with what a person in the accused’s circumstances *should* have intended. In a

---

<sup>29</sup> This presupposition fulfils an inherently moral role. The Supreme Court of Canada, for example, has indicated that the distinction between subjective and objective *mens rea* reflects differing degrees of moral fault and that the application of either standard must reflect the proportionality between this moral fault and the gravity of the offence (in accord with the principle of fundamental justice). See *Creighton* [1993] 3 SCR 3, 83 CCC (3d) 356 (S.C.C.).

world bereft of counterfactual possibility, both concepts might appear to lose their moral grounding.

Once again, however, the law makes only limited and tightly prescribed recognition of circumstances in which an agent's ability to choose is enervated to the detriment of *mens rea*. One area critical to our inquiry here is how the law treats those individuals it deems to be absolved of normal criminal liability on the basis of mental disorder.<sup>30</sup> In Canada, the special verdict of NCRMD is applied to those who at the time of the act or omission in question suffered from a mental disorder that rendered them "incapable of appreciating the nature and quality of the act or omission or of knowing that it was wrong."<sup>31</sup> Free will is not directly mentioned, yet judges, lawyers and mental health professionals often adopt the language of free will when talking about the influence of a mental disorder, something that can become all the more pronounced when neuroimaging or other seemingly reductive evidence is introduced. If such conditions are seen as sufficiently reshaping the capacity for decision-making – by distorting the defendant's view of the world so that the nature or morality of the act is no longer apparent – then there is no ontological inconsistency between the court's treatment of mentally disordered and regular defendants. If, on the other hand, the implication is that the efficacy of a condition obviates a contra-causal freedom and thereby places the defendant in a different metaphysical category, then an unwarranted distinction is made: causal determinants are at work in each case. In certain situations, the narrowing of the capacity for a range of outcomes may make it reasonable to conclude that one could not have chosen otherwise. This is as true for the "normal" individual as the "abnormal" one: a severe obsessive-compulsive may face an urge so strong that

---

<sup>30</sup> Intoxication can be another such circumstance. Despite the complexity of the current state of Canadian law regarding the defence of intoxication, the courts seem primarily concerned with intoxication as impairing the capacity to make a reasonable choice, rather than the ability to choose at all.

Given the focus here on mental disorder, it is beyond the bounds of this paper to consider the full range of legal areas where issues of choice come to bear, but the defences of necessity and duress are worth a passing consideration. Defences of necessity or duress do not technically negate *mens rea* (but instead offer a justification or excuse for forming the *mens rea* in the first place.) Nevertheless they point to a restriction of choice that is interesting to consider, in this case the result of some external force, constraint or set of circumstances that limits the reasonableness of obeying the law, requiring either a supererogatory act or, conversely, the overriding of basic human sentiment and moral imperative. In both situations, formal voluntariness is retained. The agent could have acted otherwise but an expectation that she do so is unreasonable (or unethical). The act is not unfree in the counterfactual sense of modal necessity. I could have chosen otherwise – merely at great personal risk. At an existential level, my freedom persists. At some point, however, the morally normative or instinctual pull becomes sufficiently powerful that it makes resistance genuinely impossible, perhaps even precluding its contemplation. The reality that under a libertarian or compatibilist framework there exists a tension between a presumption of freedom and a recognition that morally normative restraints on choice can bleed into areas where choice is seen as more formally constrained points to the utility of the Strawson's pragmatically grounded participant/objective distinction described below.

<sup>31</sup> See section 16, *Criminal Code*, also sections 672.1 and 672.34.

it is simply impossible to resist; a nondisordered individual may not cognitively be able to disobey the gunman's orders. As mentioned, the determination should be one of kind made with reference to the specifics of an individual's situation and condition – and not on the basis of a metaphysical switch. This pragmatic approach to freedom and unfreedom is explored in P.F. Strawson's response to the problem of free will, described below.

### A Pragmatic Approach to the Problem of Free Will

In "Freedom and Resentment," Strawson attempts to reconcile the libertarian and compatibilist positions by making what might be termed a *naturalistic* turn, moving away from purely conceptual issues about the nature of freedom and responsibility and looking at what actually goes on when we hold a person responsible.<sup>32</sup> Strawson submits that we never, actually, cease to see others as responsible beings simply as a result of accepting deterministic explanations, but only when we see them as "incapacitated" in some or all respects for ordinary inter-personal relationships."<sup>33</sup> In his view, the notions of freedom and responsibility are rooted in a "complicated web of attitudes and feelings which form an essential part of moral life as we know it."<sup>34</sup> The central strands in this web are *participant* or *reactive attitudes*: gratitude, resentment, love, hurt feelings, disgust, forgiveness, moral indignation, approbation and other normal human reactions. These attitudes and responses are a natural expression of the human concern for "whether the actions of other people...reflect...goodwill, affection or esteem, on the one hand, or contempt, indifference or malevolence on the other."<sup>35</sup>

These attitudes are modified when excusing conditions are brought to bear on the situation. Strawson distinguishes two important categories of excuse. The first aims to show that in the specific circumstances the agent lacked ill-will or disregard. ("She didn't see you there;" "He lost his balance".) Although some injury may have occurred, it was in some respect or other accidental or unintentional.<sup>36</sup> Another kind of excusing consideration goes much further than this and suggests that negative attitudes are uncalled for on the ground that the agent is somehow an "inappropriate" target because she is "abnormal or immature" and thus not fully capable of complete engagement in normal adult relationships. It is this second group that may encompass the mentally disordered or impaired. Here we adopt an *objective* stance and see the person more in terms of "social policy, as a subject for what, in a wide range of sense, might be called

---

<sup>32</sup> Peter Frederick Strawson, "Freedom and Resentment" in D Pereboom, ed, *Free Will* (Indianapolis: Hackett, 1997) 119.

<sup>33</sup> *Ibid* at 129.

<sup>34</sup> *Ibid* at 139.

<sup>35</sup> *Ibid* at 123.

<sup>36</sup> Other similar types of situation not mentioned by Strawson but where a lack ill-will mollifies resentment, would be where a degree of harm is justified by some greater good, pushing someone out of the way to save a child from an on-coming vehicle, etc.

treatment.”<sup>37</sup> But determinism does not entail that no one is capable of involvement in the moral community, that no one is “normal.” In fact, the truth or falsehood of determinism does not imply that we should universally adopt one set of attitudes over the other. It is silent about the propriety of the participant attitudes that constitute ourselves as free and responsible.

For Strawson, the belief that accepting determinism means that we should give up our participant attitudes derives from the false assumption that there exists an intelligible external standpoint from which this form of social interaction can be assessed. The relevant criteria of rationality are internal to our ways of being: “the existence of a general framework of attitudes itself is something we are given with the fact of human society. As a whole, it neither calls for, nor permits, an external ‘rational’ justification.”<sup>38</sup> This network of attitudes does not rest on any particular theoretical or metaphysical commitments beyond what is relevant to a basic human concern for the quality of attitudes manifested in interpersonal affairs.

Criticisms can be levelled at Strawson’s account from both empirical and philosophical perspectives. Although he claims to break with a conceptual and linguistic response to the free will problem and turn to the facts of “moral psychology,” any empirical investigation of those facts is markedly absent from his account. Further, Strawson admits local and temporal variation in our “concepts and practices associated with moral responsibility,” but points to “a massive central core of human thinking which has no history ... [T]here are categories and concepts which, in their most fundamental character, change not at all.”<sup>39</sup> Is there really a fundamental set of reactive attitudes seemingly rooted in human nature, or are they instead historically and culturally relative?<sup>40</sup> Relatedly, although Strawson claims “endless room” for modification, his purely descriptive metaphysics can be seen as inherently conservative,<sup>41</sup> committing him to a fixed account that blocks the generative possibility that theoretical reflection might engender more fundamental revisions to our understanding of moral responsibility.<sup>42</sup> Following this line of thought, we could also take issue with whether our reactive attitudes fully capture what it means to be a morally responsible agent. Rather than simply relying on our “innate” reactive attitudes, might not moral decision-making necessitate critical reflection on our ingrained response to see if it accords with some more dispassionate principle? Is this not in fact the essence of what it means to be a moral agent?<sup>43</sup>

---

<sup>37</sup> *Supra* note 23 at 126.

<sup>38</sup> *Ibid* at 140.

<sup>39</sup> *Ibid* at 129.

<sup>40</sup> See Michael McKenna & Paul Russell, eds, *Free Will and Reactive Attitudes: Perspectives on P.F. Strawson’s “Freedom and Resentment”* (Surrey, UK: Ashgate, 2008).

<sup>41</sup> *Ibid*.

<sup>42</sup> *Ibid* at 10-11. Also Derk Pereboom, “Alternate Possibilities and Causal Histories” (2000) 14 *Philosophical Perspectives* 119.

<sup>43</sup> Indulging the penchant of philosophers for examples that involves wiring up people’s brains can help illustrate this difficulty. McKenna and Russell ask us to imagine a case where our basic dispositions were somehow “implanted” by means of an artificial technique (e.g. neurosurgery or genetic engineering). For

These and other criticisms raise valid concerns for a Strawsonian model. Rather than adopt Strawson's program *holus bolus*, what I wish to take away is the principle of a flexible approach grounded in the practical implications of accepting or denying determinism. This not only affords a means to explain the reality that we are the products of a myriad of causal influences beyond our control while at the same time remaining moral agents, but it is also an avenue to the further development of what it means to be moral, including the appropriate way to view those mentally incapacitated. Whereas Strawson adopts a common sense descriptive approach based in a largely fixed understanding of the participant/object divide, there is no reason why this could not be expanded analytically and empirically while still retaining its practical value. Strawson advances his pragmatic course as a challenge to the epistemological basis of a problem constructed at a purely philosophical level; further developing it empirically and theoretically offers the potential to bring a prescriptive imperative rather than purely descriptive account of the balance between attitudes.<sup>44</sup>

Insight in this endeavour can be gleaned from Harry Frankfurt's account of free will, which also takes something of a naturalistic approach.<sup>45</sup> Frankfurt construes free will as the capacity to form effective second-order "volitions," reflexive desires that have as their object our first order (more basic) desires. In other words, having second-order volitions means wanting or not wanting a particular desire. Most people have this capacity to some extent, but some do not. Still others are moved contrary to or independently of their higher order desires, and as a result they lack free will. While higher order volitions are undoubtedly central to the practical exercise of free will, contrary Frankfurt, their presence is not in itself sufficient for freedom. Attitudes generated by indoctrination, brainwashing or elaborate delusion, for example, might be of a second order, but do not sit well with our notions of free will.<sup>46</sup> If a person has second-order desires but they arise from a highly paranoid view of the world, then are her actions really free? Moreover, we may identify with our second-order desires, but it is not clear from Frankfurt's arguments alone how they possess a special authority that imbues us with a potential for freedom not otherwise present.

What is missing from this type of account is the link to a more general understanding of free will within the capacity for practical reason embedded in a multi-faceted, contextual and

---

a "normal" (i.e., rational) adult who is capable of manifesting good or ill will towards others, "worries about implantation will not and cannot dislodge or discredit our reactive attitudes. Contrary to this view, however, critics will argue that implantation evidently eliminates the agent's moral responsibility and so there is something wrong with Strawson's theory" (*ibid* at 13).

<sup>44</sup> One shift that should be part of this program is away from Strawson's assumption of the exclusivity of attitudes, meaning that we demonstrate at any one time and to any particular individual either participant or objective responses. It would seem that we can and do hold both simultaneously in a varying balance. We do not (normally) temporarily cease to view the child as human merely because we realize that he is not capable of fully-realized moral decisions.

<sup>45</sup> Harry Frankfurt, "Freedom of the Will and the Concept of a Person" (1971) 68 *Philosophy* 5.

<sup>46</sup> Michael Slote, "Understanding Free Will" (1980) 77 *Journal of Philosophy* 136.

systems-based approach to the human mind. Attempts to universally distinguish the mentally “abnormal” from the “normal” by appeal to causation are wrong-headed not because distinct causal antecedents are not operative but because this in itself is much too narrow a focus to fully capture the dialectical interaction with the contextual web of circumstance (both internal and external to the mind) from which free will emerges. It is likely the case that the brain state of an individual at the precise moment of action does necessitate that action. There is a point at which the person could not have done otherwise, and this is as true for the regular individual as it is for the psychopath, psychotic or obsessive-compulsive. From a practical standpoint, however, this is a trivial, almost tautological, point. The more meaningful differentiation steps back to take in the bigger picture. Moving out from the level of micro-resolution and looking to the range of action that each person can take, we can see real and highly relevant differences emerge.

Human beings are reflexive systems, who do generate their own potential (even if they do so deterministically). At this more general level, we can talk of differing sets of possible mental states within this processual framework that pertain to each of us as individuals – ones that may vary widely in terms of range or be skewed in a particular direction (on the basis of the kinds of experiences each of us has had, our genes, manner of upbringing, etc.). At some point, various constraints may narrow or shift the range of possible outcomes to such a degree that the system either cannot generate any other possibilities and/or cannot produce ones that are socially appropriate. Returning to the neurobiological level, the existence of neurological antecedents and correlates to any conscious (or unconscious) decision is in itself unremarkable. Under a pragmatist approach, this is common to all and thus necessitates no particular change in our reactive attitudes. At a certain point, however, the structural and functional changes are such that the capacity of the person as a system is impaired to the extent that it is no longer reasonable to expect responsible or appropriate decision making.

An important component of the pragmatic approach to free will is the role of values as driving outcomes. In critiquing causal explanations of mental disorder as a criminal defence, Morse asserts that no matter how powerful his compulsion, if you put a gun at a paedophile’s head he will not molest the child.<sup>47</sup> However, if one says “molest a child in the next month and I will kill you,” he may be unable to do otherwise, even if at the level of higher order volition he wants to do anything but. Consider the irrational harm addicts can knowingly inflict upon themselves and others while seemingly unable to realize their higher order desire to quit their habit. Linking values to higher order volitions as part of a broader pragmatic approach removes some of the difficulties for the Frankfurtian account that arise from indoctrination, addiction or delusion. This capacity to reflect critically upon one’s values according to relevant criteria of practical thought and to change these values and actions through that process is central to freedom. Under this multi-faceted view, free will is not something we simply have or lack, but is an achievement, one open to various degrees of achievability.

---

<sup>47</sup> Stephen J Morse, “Craziness and Criminal Responsibility” (1999) 17 *Behavioral Sciences and the Law* at 147.

## **Conclusion: Implications of a Pragmatic Model of Freedom for the Legal System**

The implications of a pragmatic model of free will for the legal and criminal justice treatment of mental illness and impairment are manifold. Attempting a fixed description of such a system in advance is not wholly desirable, as it is something that per its own theoretical basis should be developed through praxis. A bare-bones sketch of the most fundamental repercussions is offered, however. If free will is not a binary but a spectrum, then moral agency must follow suit. The granulation of moral responsibility is a poor fit with the traditional legal binary of guilty/ not guilty (or the trinity of guilty/ not guilty/ NCRMD) and instead calls for a multi-level response that is contextualized and particularized to the specifics of an individual's circumstances. Whereas there are limits to the degree of individualization that is practical and the retention of some bright line divisions is necessary, clarifying the theoretical basis for free will and bringing the factors structuring an individual's degree of choice into rational and empirical consideration offers the possibility for a more ethical and just approach to the treatment of individuals, both "normal" and "disordered."

From an object-oriented perspective, differential treatment and intervention may be justified as appropriate and necessary. It must be based, however, on principles that balance responsibility with determinism – weighing the range of inputs that determine who we are against our capacity for choice via a value system that is itself imbricated in this process. In other words, it should be reflexive. The criteria as to when and to what degree to adopt an object-oriented perspective justifying intervention and treatment can themselves be developed as part of their own iterative engagement with the objective and participant world-view. In this way a response that recognizes the objective position need not lose sight of the ordinary humanity of the participant-based reaction.<sup>48</sup>

Beginning with the current legal approach, we can say that to the extent that the law draws broad ontological distinctions based on assumed causal differences and bases its outcomes in fixed binary categories, it is going about things the wrong way. To the extent that it looks to a pragmatic approach based in capacity it is on the right track. This practical impulse could be a starting point for a much more thorough-going development flowing from a theory of practical free will. The implication of bringing choice-structuring factors into the foreground is that a wide range of conditions, past present, and future – social context, family history, socio-economic status, community support, mental health – become relevant to the disposition of the defendant. Moral disapprobation is not absent from this system but it is more fluid. At one end of the spectrum, a finding of not guilty would continue to mean that no criminal legal consequence whatsoever is attached. On the other hand, conditions where an individual had little constraints

---

<sup>48</sup> And it should always be recognized that there likely are epistemological limits on the extent to which fundamental properties of persons can be ascertained via an object-oriented attitude – the degree to which can the self contain itself as an object of its own knowledge. As well, the autopoietic quality of mind – the capacity for self-creation that emerges through dialectical engagement with self – does not have to be outside of physical law to unfold in non-algorithmic and not fully predictable ways.

on his ability to choose freely could accrue significant legal culpability and moral stigma. In other circumstances, an accused for whom factors sufficiently limited his ability to choose (because of neurocognitive impairment, genetic disposition, physical abuse or the combination of all) would see a finding that he was not free to have acted otherwise, even if today he might fail to meet the current NCRMD standard (as he was to some degree aware of nature and morality of his actions).

The range of dispositions would be much more varied, responding to the same multifactorial variables that act to limit the defendant's circumstances. This would require the integration of disparate ministries and agencies, not just the penal system (whose presence could likely be significantly reduced) but health, housing and other social service agencies. While (from the objective pole) utilitarian ends could still be served, in a general sense I believe the system would look much more restorative and much less retributive than it does at present. In this way it would seek to reconcile an empirical understanding of the conditions and experiences that make us who we are and the reality of human agency, thereby increasing the dignity, compassion and effectiveness of the criminal justice response.