


2004

# Reason, Results, and Criminal Responsibility

Stephen J. Morse

*University of Pennsylvania*, [smorse@law.upenn.edu](mailto:smorse@law.upenn.edu)

Follow this and additional works at: [http://scholarship.law.upenn.edu/faculty\\_scholarship](http://scholarship.law.upenn.edu/faculty_scholarship)

 Part of the [Criminal Law Commons](#), [Ethics and Political Philosophy Commons](#), [Law Enforcement and Corrections Commons](#), and the [Metaphysics Commons](#)

---

## Recommended Citation

Morse, Stephen J., "Reason, Results, and Criminal Responsibility" (2004). *Faculty Scholarship*. Paper 527.  
[http://scholarship.law.upenn.edu/faculty\\_scholarship/527](http://scholarship.law.upenn.edu/faculty_scholarship/527)

This Article is brought to you for free and open access by Penn Law: Legal Scholarship Repository. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of Penn Law: Legal Scholarship Repository. For more information, please contact [PennlawIR@law.upenn.edu](mailto:PennlawIR@law.upenn.edu).

# REASON, RESULTS, AND CRIMINAL RESPONSIBILITY<sup>†</sup>

Stephen J. Morse\*

*In this article, Professor Morse provides new insights into the concept of desert in criminal punishment. Professor Morse argues that intentional action and forbearance are the only kinds of human conduct that can be effectively guided by the criminal law. The consequences of action, however, cannot be fully guided and are therefore inappropriate predicates for desert. Professor Morse contends that a rational system of criminal law should focus solely on actions and should not impose punishment based on results.*

*Professor Morse's action-guiding account of the law helps to explain disputed areas of criminal law, including attempt liability, risk creation, causation, accomplice liability, strict liability, and the justifications. After responding to the counterarguments of leading criminal law scholars, the article concludes that a consistent subjectivism concerning criminal liability is both possible and fair.*

*'As you can plainly see, failed guidance  
is the cause the world is steeped in vice,  
and not your inner nature that has grown corrupt.*

—Dante\*\*

## TABLE OF CONTENTS

I.	Introduction .....	364
II.	Persons, Reasons, and Responsibility .....	367

<sup>†</sup> (c)2003

\* Ferdinand Wakeman Hubbell Professor of Law & Professor of Psychology and Law in Psychiatry, University of Pennsylvania Law School.

This article was presented at faculty workshops at Chicago, Fordham, N.Y.U., Seton Hall, Texas, USC, and Virginia, and at a faculty retreat at Penn. Thanks to Barry Adler, Scott Altman, Larry Alexander, Jody Armour, Ted Blumoff, Sherry Colb, John Gardner, Ron Gare, David Garland, Ed Greenlee, Kyron Huigens, Heidi Hurd, Jim Jacobs, Leo Katz, Kimberly Kessler Ferzan, Paul Litton, Michael Moore, Stephen Perry, David Richards, Paul Robinson, David Rudovsky, Larry Sager, Stephen Schulhofer, and Dan Simon for their gracious assistance, and to the members of the audience at the venues in which this article was presented for helpful comments. As always, my personal attorney, Jean Avnet Morse, furnished sound, sober counsel and moral support.

\*\* DANTE ALIGHIERI, PURGATORIO XVI:103–105 (Jean Hollander & Robert Hollander trans., 2003).

364	UNIVERSITY OF ILLINOIS LAW REVIEW	[Vol. 2004
III.	Results and Luck .....	378
IV.	Doctrinal Applications .....	386
A.	Attempt Liability.....	387
1.	Last-Act, Completion, and Preparation.....	387
2.	Impossible Attempts .....	391
3.	Inherent Impossibility .....	392
4.	Renunciation or Abandonment .....	392
B.	Risk Creation.....	393
C.	Causation.....	395
D.	Accomplice Liability.....	397
E.	Strict Liability .....	400
F.	Justification .....	405
G.	Results and Epistemic Value Added .....	409
V.	The Counter-Argument: Results Matter .....	412
A.	Michael Moore on the Independent Relation of Results to Moral Desert.....	413
1.	Moral Coherentism and Results .....	415
2.	Emotions, Experience, and Desert .....	418
B.	Leo Katz and Trying Transitivity .....	422
C.	Stephen Perry and a Different Form of Blameworthiness .....	425
D.	Other Arguments for Why Results Contribute to Desert .....	426
E.	Psychological or Sociological Necessity.....	427
VI.	Conclusion: Consistent Subjectivism.....	428
VII.	Appendix: The Metaphysical Challenge to Responsibility .....	430
A.	Incompatibilism .....	431
1.	Hard Determinism.....	432
2.	Libertarianism .....	433
B.	Compatibilism.....	437

## I. INTRODUCTION

The criminal law creates and reflects value by announcing which conduct is sufficiently wrong to deserve blame and punishment. It guides conduct by giving citizens good reason to comply, both by manifesting the underlying moral justification for the law and by providing incentives to obey.<sup>1</sup> This article is a contribution to the theory of desert, and assumes, with most criminal law theory, that desert is at least a necessary condition of just punishment.<sup>2</sup> It also assumes that theories and practices

---

1. In the case of most crimes that vitally concern us—the core crimes against the person and property—the agent’s conduct is morally wrong. I recognize that much criminal law is now “regulatory” and does not address conduct most people think is morally wrong. Nonetheless, if a criminal prohibition is justifiable, it is *prima facie* good to obey the law.

2. There are many types of desert theory, but all have in common a retrospective evaluation of an agent’s conduct to determine if that conduct meets desert criteria. If so, desert justifies an appropriate, proportionate legal response to that conduct, such as blame and punishment, for no reason other than that the agent genuinely deserves the response. Not all desert theories claim that the law

of responsibility, blame, and punishment are, in principle, coherent and that a just criminal law is therefore possible. Thus, I take the perspective of an internal but critical participant<sup>3</sup> in our criminal law and attempt to make it more rational and fair.

This article suggests that intentional action and forbearance are the only aspects of human conduct that potentially can be fully guided consciously, explicitly, and effectively by moral or legal rules. Although this claim is hardly novel, when properly understood it has profound consequences for the criminal law. Most generally, although criminal prohibitions and adjudication are in part evaluative, focusing on the action-guiding function of criminal law suggests that intentional action and forbearance are the only proper objects for moral evaluation and the ascription of moral and criminal desert. Imposing blame and punishment for anything other than intentional action and forbearance is both unfair and, in most cases, useless. More specifically, understanding the centrality of the action-guiding function clarifies and perhaps solves many traditional puzzles of criminal liability, such as the relation of results to desert and the justification of justifications, and serves as a guide to rational reform of doctrines, such as those governing attempt liability and risk creation.

Part II is theoretical and sets forth the argument for the criminal law's primary action-guiding function. I begin by considering the concept of the person as a being with the capacity for practical reason and the law's view of the person as a practical reasoner: an agent who acts for reasons. I claim that the law's action-guiding function is inherent in the nature of law itself and that this function explains our concepts of responsibility and the excusing conditions. This part also explains why the criminal law does not and should not punish for intentions alone, even if

---

must give an agent what he or she deserves for past conduct, but all hold that the agent does deserve the response if the agent's conduct meets the criteria for desert.

Experimental evidence suggests that human beings have evolved to engage in nonutilitarian, "altruistic" punishment when members of their social group "defect" by violating a normative expectation. See Ernst Fehr & Simon Gächter, *Altruistic Punishment in Humans*, 415 NATURE 137, 137 (2002). Other cooperative primates exhibit behavior best interpreted as expressing resentment when they perceive that they are treated undeservedly unequally. See Sarah F. Brosnan & Frans B. M. de Waal, *Monkeys Reject Unequal Pay*, 425 NATURE 297, 297 (2003). In recent simulated jury studies, Cass Sunstein and associates have demonstrated that people from every demographic group are uniformly "intuitive retributivists" who reject common and central theses of economic and utilitarian theory, such as a preference for optimal deterrence. Cass R. Sunstein, *Outrage 2* (Oct. 6, 2002) (unpublished manuscript, on file with author).

I freely concede that no uncontroversial justification for state punishment exists, and I certainly cannot provide one. Many would agree, however, that plausible accounts of both retributive and consequential justifications have been given, and it is clearly the case that our system of criminal justice assumes that people should not be punished unless punishment is deserved.

Although this article is a contribution to the theory of desert, it is also consistent with consequentialist approaches to criminal liability. When inconsistency arises, consequential concerns usually must yield in a system in which desert is a necessary condition of just punishment. At the least, a compellingly strong consequentialist case must be made for ignoring desert.

3. I borrow this apt characterization, with slight amendment, from Thomas Merrill. Thomas W. Merrill, *The Landscape of Constitutional Property*, 86 VA. L. REV. 885, 890 (2000).

intentions are the royal road to culpability assessment. Most generally, I explain why intentional action and forbearance are the only aspects of human conduct that can be guided by criminal law and are the only proper objects for ascriptions of criminal responsibility and desert. In brief, the law cannot directly command or guide states of affairs, but it can guide good action, which in turn can create states of affairs.

Part III examines in detail an issue raised briefly in part II: What is the contribution of “result luck” to fair ascriptions of desert and the consequent imposition of blame and punishment? I demonstrate that many of the arguments concerning luck that dominate discussion of this issue are misguided because they are really about the broader question of whether responsibility is possible in a deterministic world. I argue that “moral luck” poses no genuine problem for the criminal law within an action-guiding model of desert, and that a plausible metaphysics, compatibilism, furnishes an answer to those who think that all behavior is not “up to us” and that responsibility is impossible. I conclude that the consequences of action cannot be fully guided and are thus not appropriate predicates for desert. Full culpability and desert are established by intentional action that risks a harmful result.

The next part addresses exemplary doctrinal implications of the action-guiding approach, which considers practical reason to be the touchstone of responsibility and desert. The substantive contexts addressed are attempt liability, risk creation, causation, accomplice liability, strict liability, and justifications. Although the implications of the action-guiding account of these doctrines are intrinsically interesting and imply substantial reforms of current criminal law, which I discuss, they are included primarily as examples of the virtues of the approach. Throughout, I argue that the subjectivist, *ex ante* action-guiding account provides the most satisfactory basis for desert in all these areas and serves as a guide to rational reform. Part IV concludes with a brief examination of the evidentiary value of results evidence.

Part V canvasses and rejects various counterarguments that results (and therefore causation) do contribute to an offender’s desert. It considers the leading arguments from criminal law scholars and general alternative arguments that results contribute to desert.

Part VI is a brief conclusion that suggests that consistent subjectivism concerning criminal liability is possible and fair. The appendix addresses in detail the fundamental metaphysical challenge, discussed briefly in part III, that the apparently causal nature of the universe poses to the concepts of moral responsibility and desert inherent in the action-guiding account. I place this material in an appendix because the article adopts an internal participant perspective, and the challenge from determinism is external. I conclude that compatibilism—the view that moral responsibility is consistent with determinism, universal causation, and like metaphysical accounts of the physical universe—is the best the-

ory to explain our practices, and the only theory that can plausibly justify them.

## II. PERSONS, REASONS, AND RESPONSIBILITY

The criminal law is an intensely practical enterprise that seeks to prevent culpably produced harmful states of affairs, because few things are as important to a good life, to human flourishing, as freedom from the malevolent behavior of others. Criminal prohibitions give agents good moral reasons not to offend by announcing which conduct is wrong, and good prudential reasons by threatening to punish and then punishing those who offend the law. I make the simplifying assumptions that most malum in se crimes describe uncontroversially wrongful behavior unless a justification is claimed, and that most citizens fully understand why engaging in such behavior would be wrong even in the absence of a formal criminal prohibition. In brief, a crucial aim of the criminal law is to guide and thus to prevent certain *actions*, because only actions can culpably produce harms.

Human action is distinguished from all other phenomena because only action is explained by reasons resulting from desires and beliefs, rather than simply by mechanistic causes. Only human beings are fully intentional creatures. To ask why a person acted a certain way is to ask for reasons for action, not for reductionist biophysical, psychological, or sociological explanations.<sup>4</sup> Practical reason is inescapable for creatures like ourselves who inevitably care about the ends they pursue and about what reason they have to act in one way rather than another.<sup>5</sup> Only people can deliberate about what action to perform and can determine their conduct by practical reason.<sup>6</sup>

The law clearly treats people as intentional agents and not simply as part of the biophysical flotsam and jetsam of the causal universe. It could not be otherwise. To employ a useful oversimplification, law and morality are systems of rules that at the least are meant to guide or influence behavior and thus to operate as a potential cause of behavior.<sup>7</sup> But

---

4. To comprehend fully why an agent has particular desires, beliefs, and reasons requires biophysical, psychological, and sociological explanations. Ultimately, however, human action is more than simply the mechanistic outcome of mechanistic variables.

5. HILARY BOK, *FREEDOM AND RESPONSIBILITY* 75–91, 128–31, 146–51 (1998).

6. In a personal communication, David Garland has noted that we sometimes speak of organizations deliberating. Although we talk metaphorically about this phenomenon, I believe it is just a metaphor. Nonetheless, if one were convinced that groups can be treated as morally equivalent to individual agents, I see no obvious reason why the criminal law implications of the action-guiding account should not apply to entities such as corporations.

7. See John R. Searle, *End of the Revolution*, N.Y. REV. BOOKS, Feb. 28, 2002, at 35. Searle writes:

Once we have the possibility of explaining particular forms of human behavior as following rules, we have a very rich explanatory apparatus that differs dramatically from the explanatory apparatus of the natural sciences. When we say we are following rules, we are accepting the notion of

legal and moral rules are not simply mechanistic causes that produce “re-flex” compliance. They operate within the domain of practical reason. Agents are meant to and can only use these rules as potential *reasons for action* as they deliberate about what they should do. Moral and legal rules thus guide actions primarily because they provide an agent with good moral or prudential reasons for forbearance or action. Unless people were capable of understanding and then using legal rules as premises in deliberation, law would be powerless to affect human behavior.<sup>8</sup> I am not suggesting that human behavior cannot be modified by means other than influencing deliberation or that human beings always deliberate before they act. Of course it can and of course they do not. But law operates through practical reason, even when we most habitually follow the legal rules. Law can directly and indirectly affect the world we inhabit only by its influence on practical reason. I believe that this view is consistent with virtually any jurisprudential theory about the essential nature of law and that such consistency is an attractive feature of the view.

All things being equal, intentional action or forbearance is the only aspect of the human condition that is fully “up to us,” that is fully within our control, and that can be fully guided by and produced by our reason.<sup>9</sup> In the entire chain of causation that leads to compliance with or breach of a moral or legal obligation, only action is potentially fully determined

---

mental causation and the attendant notions of rationality and existence of norms. . . . The content of the rule does not just describe what is happening, but plays a part in *making it happen*. *Id.*; see also LARRY ALEXANDER & EMILY SHERWIN, *THE RULE OF RULES: MORALITY, RULES & THE DILEMMAS OF LAW* 11–25 (2001) (explaining why rules are necessary in a complex society and contrasting their account with H. L. A. Hart’s theory). My account also applies to noncodified norms, which likewise operate as potential reasons for action.

8. See Scott J. Shapiro, *Law, Morality and the Guidance of Conduct*, 6 *LEGAL THEORY* 127, 131 (2000). This view assumes that laws (and morality) are sufficiently knowable to guide conduct. A contrary assumption, however, is largely incoherent. As Shapiro writes:

Legal skepticism is an absurd doctrine. It is absurd because the law cannot be the sort of thing that is unknowable. If a system of norms were unknowable, then that system would not be a legal system. One important reason why the law must be knowable is that its function is to guide conduct.

*Id.* I do not assume that legal rules are always clear and thus capable of precise action guidance. If most rules in a legal system were not sufficiently clear most of the time, however, the system could not function. Moreover, the rules of the criminal law tend to be particularly clear because the U.S. Constitution prohibits punishment in the absence of clear notice.

I recognize that ubiquitous criminal law standards, such as reasonableness, seem to present a problem, but the common law interpretive process has apparently been adequate to avoid potential constitutional pitfalls. See generally Jeremy Horder, *Criminal Law and Legal Positivism*, 8 *LEGAL THEORY* 221, 233–37 (2002). Professor Horder argues that if the criminal law’s criteria for mala in se crimes are to express moral considerations, as they should, it is impossible and unwise for the legislature to draft utterly explicit, objective criteria. Open-ended or allusive terms, such as “reasonableness,” will thus be inevitable, but will provide limited ex ante guidance to citizens. They can, however, provide ex post guidance to judges and juries. Although Horder is correct that fully objective ex ante guidance is a chimera, more allusive, moral criteria, as interpreted by the courts, surely provide sufficient guidance to satisfy the principle of legality.

This article does not address jurisprudential questions concerning law’s authority generally, or whether legal rules are binding only if they comport with morality.

9. For the sake of simplicity, I will henceforth refer only to the case of action. But the analysis applies equally to the case of forbearance.

by reasons, even if an agent does not deliberate about a specific action. The universe's causal laws, our causal backgrounds, our opportunity sets, and the events of the universe other than our actions themselves are not up to us, not dependably the potential product of reason. Although we can intentionally produce specific thoughts, feelings, and emotions, much cognition, affect, and emotion is not intentional and simply happens to us. Thoughts pop into our heads, seemingly for no reason; feelings come upon us unbidden. We can use reason to try to modify our character—those traits and dispositions that seem to motivate or otherwise cause intentional action—but this is notoriously hard to do. At most we can nibble at the margins. Thoughts, desires, and character are not primarily a product of our reason. We can of course use reason to guide us toward benign environments and away from those that would tempt us to evil, but even the greatest caution and foresight will not prevent an appointment in Samarra.<sup>10</sup> Opportunity is not primarily a product of reason.

We can use reason to predict the consequences of various actions and thus to decide which to perform. In both a metaphorical and partially literal sense, then, reason can guide the consequences of our actions. But even the greatest degree of knowledge and attention will not prevent the forces of nature, including other human actions, from sometimes intervening to prevent (or to complete in unpredictable ways) the predictable, natural, and probable consequences of our actions (or non-actions). Our reason plays no role in those forces and could do so only if each of us had perfect knowledge of and perfect power over all the causal variables of the universe.

The claim that reason can guide only human action does not imply that human beings are always guided by reason, including the reasons provided by moral and legal rules. Much human action is unreasonable, thoughtless, foolish, irrational, arational, and the like. Nevertheless, the capacity to guide one's actions by reason develops through childhood and adolescence and is present in most adults. Successful human interaction, not to mention the survival of the species, would be impossible without this capacity. It is the touchstone of moral and legal responsibility. Moreover, when important rights and interests are at stake, we expect each other to use our capacity to be guided by reason, once again including reasons the law provides, to avoid risking harm to ourselves or others. In addition, the criminal law guides socializing agents, such as families, schools, and religious organizations, as they attempt to inculcate law-abiding attitudes, habits, and behaviors. The failure of morality and law always to guide action successfully does not undermine the claim that the action-guiding function is crucial to the theoretical and practical importance of law and morality in human interaction.

---

10. See <http://www.ie.ispace.org/books/apf/the-colour-of-magic.html> (last visited Sept. 16, 2003) (discussing the origins of the allegory of Samarra).



The preceding discussion, which emphasizes responsibility for action, has assumed that the definition of "action" is clear, but the claim that only action can fully be guided by reason, including the reasons morality and law provide, requires some account of action. After all, if we cannot identify actions, then the action-guiding approach (and much of moral and legal evaluation of responsibility more generally) collapses.

Most uncontroversially, an action is an intentional bodily movement, that is, a bodily movement done for some reason, or at least rationalizable or guidable by a reason.<sup>11</sup> Such general definitions tell us very little, however. Indeed, it is famously the case in the philosophy of mind and action that many descriptions can accurately describe what an agent did and that the metaphysics of action are insufficiently understood to provide an uncontroversial method of describing and individuating actions.<sup>12</sup> For example, how should we describe an agent's (*A*)causing the killing of another by shooting the victim (*V*) through the heart at point blank range with the intent to kill the victim? What is the correct *act* description? Consider the following possibilities: *A* killed *V*; *A* intentionally killed *V*; *A* made *V*'s children orphans (assuming that the other parent was already dead); *A* intentionally shot *V* through the heart with the intent to kill *V*; *A* intentionally shot at *V*'s heart with the intent to kill *V*; *A* intentionally aimed the gun at *V*'s heart and intentionally pulled the gun's trigger with the intent to shoot *V* through the heart and with the intent to kill *V*. All, I would submit, are perfectly accurate descriptions of what *A* "did" and many more such descriptions could be generated. Neither philosophy nor psychology can provide an uncontroversial resolution to deciding which description to choose, and the criminal law can neither conceptually solve those problems nor wait for others to do so. For criminal law, then, the question is which description most accords with the criminal law's purposes without being conceptually and practically unworkable.

Legal definitions of crimes do not express a monolithic conception of action and contain many different types of (culpably produced) act descriptions. The following examples are just a sample. The act requirement of some crimes is defined in terms of results that are at least tempo-

---

11. See, e.g., Donald Davidson, *Actions, Reasons, and Causes*, in *ACTIONS & EVENTS* 3 (2d ed. 2001). The qualification that an action is at least potentially rationalizable or guidable by reason addresses cases like habitual actions, which are clearly intentional but may seem entirely "thoughtless." I do not address in this article normative controversies concerning the criteria for "good" reasons. This is an immense topic within the theory of action and rationality that is beyond my present purposes. For my account, it is sufficient that people have reasons that causally explain their behavior.

By suggesting that reasons can causally explain behavior, I am not adopting a dualist position concerning the mind-body problem. I assume that this process depends on physically explicable brain activity. See *infra* note 20 and accompanying text.

I recognize that an agent can intentionally engage in mental activity, cause herself to experience a feeling, and the like. Thus, such purely psychological "doings" could be considered acts. I address this question *infra* Part II.

12. See, e.g., ROBERT AUDI, *ACTION, INTENTION AND REASON* 1-4 (1993) (describing the "basic philosophical divisions" in each of the four major problem areas in action theory).

rally distinct from the intentional bodily movements that cause those results. The homicide prohibition against (culpably) killing is the most obvious example. The act requirements of other crimes appear to prohibit certain (culpably produced) specific bodily movements without regard to further results. For example, perjury is committed if a witness (culpably) utters a falsehood about a material fact while under oath, whether or not anyone believes the falsehood. For another example, larceny is committed if an agent (culpably) takes and carries away the property of another, whether or not the owner notices or ultimately loses possession of the property. In all these cases, the act definition contains no highly specific bodily movement description, and many different types of bodily movements are sufficient to satisfy the criminal law's act definition. For instance, killing could be done by shooting, poisoning, knifing, bludgeoning, strangling, scaring, and a host of other bodily movements. Uttering a material falsehood about the same issue could be accomplished by using many different combinations of words. Property can be taken and carried away by use of one's unaided body or by a forklift, and so on.

What is required for culpability in every crime, however, is that the agent must intentionally perform some bodily movement that is sufficient to accomplish the prohibited act.<sup>13</sup> Unless such a sufficient act is performed, no liability for completed crime—the ultimately undesirable state of affairs the law seeks to avoid—is possible. Properly understood, then, the law prohibits culpably engaging in conduct sufficient to produce the prohibited, harmful state of affairs in the ordinary course of events. This follows from the guidance perspective because the reasons to cause or not to cause the prohibited state of affairs, such as killing or lying under oath, are precisely the same reasons to perform or not to perform the actions, such as killing conduct or uttering falsehoods, that would cause those states in the ordinary course of events. Reason can surely guide the intention to engage in killing and lying conduct, but once the sufficient act is performed, reason cannot guide what occurs thereafter.

For criminal law purposes, then, an act is any intentional movement accompanied by a culpable mental state. Reason can guide the agent until the act occurs, but once it occurs, it is done and cannot be taken back. At that point, its consequences cannot be guided by reason. Again, luck of endowment, character, and opportunity may increase or decrease the probability that such an act will be performed, and luck concerning the “response” of the physical world to an agent's action may increase or decrease the probability that the harmful state of affairs will occur. Among normal agents, however, only intentional action is fully guidable by reason.

This account faces a number of related challenges. First, it fails to address the possibility of mental acts. Not all thoughts simply pop into

---

13. Intentionally uttering words—so called speech acts—is surely included. For example, the act required by the definition of solicitation and conspiracy is accomplished primarily by words.

an agent's head; some are the product of intentional mental effort. Suppose an agent were asked to add twenty-five and thirteen in the agent's head. Determining the sum mentally and silently would surely be an intentional act, but there would be no bodily movement. Thus, the definition of action could be expanded beyond cases of intentional bodily movements performed for a reason, and mental acts are plausible objects of moral evaluation according to some moral theories. For example, an agent who intentionally brings to mind scenarios involving the infliction of suffering because such scenarios cause the agent great pleasure arguably suffers from a moral defect. Recognizing that the definition may be thus expanded poses little problem for the law, however, because the harm principle suggests that intentional mental acts wrong no one other than the agent himself or herself, even if they are immoral.<sup>14</sup> Such mental acts should not be criminalized in a society that rejects the use of the criminal law simply to enforce human virtue, even assuming that the law were capable of detecting and proving such immoral mental acts.

A potentially more serious challenge is a regress argument, the possibility that the genuine ground for culpability and desert is mental acts and that the bodily movement/act requirement in criminal law should be abandoned. Action has been defined as intentional bodily movements rationalizable by reasons. Despite scientific and philosophical advances, understanding action is still largely a philosophical and scientific mystery.<sup>15</sup> Roughly speaking, materialist and folk psychological theories compete with little resolution in sight—at least until the mind-body prob-

---

14. See generally JOEL FEINBERG, HARM TO OTHERS 6–64 (1984) (defining harm, and describing the necessity for a harm principle in a liberal state).

15. SAMUEL GUTTENPLAN, MIND'S LANDSCAPE: AN INTRODUCTION TO THE PHILOSOPHY OF MIND 169–88 (2000) (describing different approaches toward explaining human action in terms of the interaction between brains and reasons); G. F. Schueler, *Action Explanations: Causes and Purposes*, in INTENTIONS AND INTENTIONALITY: FOUNDATIONS OF SOCIAL COGNITION 251, 252–57 (Bertram F. Malle et al. eds., 2001). Or, as Thomas Nagel writes: “There seems no room for agency in a world of neural impulses, chemical reactions, and bone and muscle movements. Even if we add sensations, perceptions, and feelings we don't get action, or doing—there is only what happens.” THOMAS NAGEL, THE VIEW FROM NOWHERE 111 (1986). Ian McEwan provides a brilliant literary account that captures the mystery:

Briony sat on the floor. . . . She raised one hand and flexed its fingers and wondered . . . how this thing, this machine for gripping, this fleshy spider on the end of her arm, came to be hers, entirely at her command. Or did it have some little life of its own? She bent her finger and straightened it. The mystery was in the instant before it moved, the dividing moment between not moving and moving, when her intention took effect. It was like a wave breaking. If she could only find herself at the crest, she thought, she might find the secret of herself, that part of her that was really in charge. She brought her forefinger closer to her face and stared at it, urging it to move. It remained still because she was pretending, she was not entirely serious, and because willing it to move, or being about to move it, was not the same as actually moving it. And when she did crook it finally, the action seemed to start in the finger itself, not in some part of her mind. When did it know to move, when did she know to move it? There was no catching herself out. It was either-or. There was no stitching, no seam, and yet she knew that behind the smooth continuous fabric was the real self . . . which took the decision to cease pretending, and gave the final command.

IAN MCEWAN, ATONEMENT 33–34 (2001). No one really knows how Briony's finger was crooked.

lem is solved.<sup>16</sup> There are also major disputes within each of the camps.<sup>17</sup> Because the law employs the folk psychological model of human action, however, I shall address the regress argument in those terms.

The general picture of action in folk psychology is the practical reason view. Simply put, agents have desires and beliefs, and form intentions to act so as to satisfy those desires in light of their beliefs.<sup>18</sup> Anyone reading this article had a desire to do so (for some reason) and believed that reading it would satisfy that desire. Thus, the reader formed the intention to read the article, which is a mental state. The perplexing question is how the intention became executed by the bodily movements of picking up, opening, and reading the article. The short answer is that no one knows how this happens, but the best theory is that agents possess volitional abilities—sometimes referred to as the “will”—that execute those intentions by causing the sufficient bodily movements.<sup>19</sup> On this account, volition is another type of intention and is thus also a mental state. There can be no action without a volition, and once the volition is formed, the action will occur. The body seemingly just does what the mind tells it to do.<sup>20</sup> But if so, the volition is what the agent actually “does,” and the bodily movement is just a biophysical consequence. To see this, imagine—hypothetically to be sure—that at the *precise* moment that an agent’s body was about to move in response to a volition, say, to strike another unjustifiably, one of the following occurred: a much stronger person grabbed the potential striker’s arm and entirely immobilized it, or a concurrent neurophysiological event paralyzed the relevant muscles. In a very real sense, the agent did everything she could do to produce a prohibited bodily movement, a prohibited action, and it was just a matter of luck, unguided by reason, that prevented the action from occurring. It seems that volitions, and not actions, are what can finally be guided by reason.

---

16. The scientific, moral, and practical consequences of solving the mind-body problem, if it can be solved, are likely to be revolutionary and incalculable. See PAUL R. MCHUGH & PHILLIP R. SLAVNEY, *THE PERSPECTIVES OF PSYCHIATRY* 11–12 (2d ed. 1998).

17. See PAUL M. CHURCHLAND, *MATTER AND CONSCIOUSNESS* 45–49 (3d prtg. 1984) (comparing eliminative materialism to folk psychology).

18. The proper description of an adequate folk psychological explanation is itself controversial, but the simplification in the text is adequate for our purposes. Compare Alfred R. Mele, *Acting Intentionally: Probing Folk Notions*, in *INTENTIONS AND INTENTIONALITY: FOUNDATIONS OF SOCIAL COGNITION*, *supra* note 15, at 27 and Schueler, *supra* note 15, at 257–63, with Bertram F. Malle, *Folk Explanations of Intentional Actions*, in *INTENTIONS AND INTENTIONALITY: FOUNDATIONS OF SOCIAL COGNITION*, *supra* note 15, at 265.

19. See MICHAEL S. MOORE, *ACT AND CRIME: THE PHILOSOPHY OF ACTION AND ITS IMPLICATIONS FOR CRIMINAL LAW* 113–65 (1993) (giving the most detailed account within the legal literature of the nature of volition).

20. I recognize that this locution sounds alarmingly dualist—as if I accept a Cartesian view that we have minds independent of our bodies that can interact with our bodies and somehow cause them to move. I am a thoroughgoing “matter first” materialist, however, and use such locutions because it is almost impossible to describe action otherwise. See generally JANET RADCLIFFE RICHARDS, *HUMAN NATURE AFTER DARWIN: A PHILOSOPHICAL INTRODUCTION* 51–66 (2000) (comparing mind-first to matter-first approaches to human behavior).

The volitional story is not implausible as a philosophical account of irreducible human “doing,”<sup>21</sup> but it is ultimately unconvincing as a challenge to the act requirement in criminal law. To begin, the regress argument proves too much. Any necessary precondition for any bodily movement can always in principle be blocked or deflected by various events. A volition is presumably the agent’s necessary and sufficient last mental act prior to performing an intentional bodily movement, but the desires and beliefs that produce the volition can also be blocked or deflected. This regress potential would obliterate all responsibility, a possibility I discuss below.<sup>22</sup> Assuming that we do not abandon responsibility evaluation and practices, the most telling counterargument is intensely practical. In the ordinary course of human affairs—indeed, virtually always—volitions of neurologically intact agents are followed instantaneously by action.<sup>23</sup> If a volition is formed, in almost all cases the mental becomes the physical. Exceptions are so rare, and so difficult to identify, that they may be entirely theoretically ignored with no substantial loss to the justice of the criminal law.

All of the foregoing considerations concerning practical reason, action, and the nature of the criminal law explain, finally, why desert depends ultimately on action and not on thoughts, desires, character, opportunities, and outcomes; and why we are not fully responsible for our thoughts, desires, characters, and opportunities.<sup>24</sup> What we do is influenced by the thoughts, desires, characters, and opportunities that we have or sometimes intentionally create, but only intentional action can

---

21. Many philosophers would claim that all action is or includes a “trying.” Severin Schroeder, *The Concept of Trying*, 24 PHIL. INVESTIGATIONS 213, 213 (2001) (calling this claim the “ubiquity thesis”).

22. See *infra* Part VII.

23. Some might argue that the volition and the action are conceptually indistinguishable. See NAGEL, *supra* note 15, at 111 (1986) (“[T]he only solution is to regard action as a basic mental or more accurately psychophysical category—reducible neither to physical nor to other mental terms.”). Conceptually, however, a temporal sequence does not seem implausible, and there is suggestive empirical evidence to support the sequential assumption. Benjamin Libet, *Do We Have Free Will?*, in *THE VOLITIONAL BRAIN: TOWARDS A NEUROSCIENCE OF FREE WILL* 47, 51 (Benjamin Libet et al. eds., 1999) (describing empirical research demonstrating that what is termed the “conscious will” precedes muscle activation by about 150 milliseconds); Benjamin Libet, *Do We have Free Will?*, in *THE OXFORD HANDBOOK OF FREE WILL* 551, 551–60 (Robert Kane ed., 2002) (same). But see HENRIK WALTER, *NEUROPHILOSOPHY OF FREE WILL: FROM LIBERTARIAN ILLUSIONS TO A CONCEPT OF NATURAL AUTONOMY* 250–52 (Cynthia Klohr trans., 2001) (criticizing Libet’s interpretation of the experimental findings); Jing Zhu, *Reclaiming Volition: An Alternative Interpretation of Libet’s Experiment*, 10 J. CONSCIOUSNESS STUD. 61, 67–74 (2003) (providing an alternative interpretation of Libet’s results based on artifacts in Libet’s experimental design and on the results of experiments by other investigators).

Even assuming the validity of Libet’s fascinating findings does not cast doubt on the existence of behavioral causation by intentions. Libet simply demonstrates that nonconscious brain events precede conscious experience. This seems precisely what one would expect of the mind-brain, and does not mean that the intentionality played no causal role. Libet also concedes that people can “veto” the act, which is another form of mental act that plays a causal role. Libet’s work is fascinating, but it does not prove that persons are not conscious, intentional agents.

24. See *infra* Part III (arguing in detail that results should not affect desert for many of the same reasons).

be guided by considerations that give us reason to respond to our thoughts, desires, characters, and opportunities in ways that are possible for us. If genuine responsibility and desert are possible at all, then only intentional action can form their basis.

This explanation of responsibility and desert also explains the excusing conditions. The primary, generic excusing condition is lack of the capacity for rationality, which is precisely the inability to be guided by good reason in a particular context. Infancy, cognitive tests for legal insanity, and partial excusing conditions are classic examples.<sup>25</sup> The inability to be guided by good reason in the situation also explains so-called compulsion excuses better than explanations based on mechanistic uncontrollability.<sup>26</sup> For example, the test for duress asks whether a “person of reasonable firmness” would yield to a coercive threat and act so as to produce a negative balance of evils.<sup>27</sup> I believe that the best interpretation of this test is that some threats produce such a hard choice that a person simply cannot be guided by the right in those circumstances. Explanations of the excuses other than irrationality and a limited form of compulsion have been offered, but virtually all are confused, fail to explain our practices, or prove too much.<sup>28</sup>

The action-guiding account also explains perplexing problems about excuses and how they should be approached. Consider, for example, whether psychopaths should be held responsible for their criminal deeds. As traditionally diagnosed, the psychopath is an egocentric agent who utterly lacks empathy and conscience.<sup>29</sup> For most people, of course, the capacity for empathy and the internal moral compass of conscience furnish the most compelling reasons not to breach moral and legal expectations. Indeed, if the law’s threats were the primary prophylactic against crimi-

---

25. See, e.g., MODEL PENAL CODE § 4.01(1) (1962) (providing the legal insanity test); *id.* § 210.3(b) (holding that “extreme mental or emotional disturbance” reduces a killing that would otherwise be murder to manslaughter).

26. See Stephen J. Morse, *Crazy Reasons*, 10 J. CONTEMP. LEGAL ISSUES 189, 204 (1999); Stephen J. Morse, *Hooked on Hype*, 19 LAW & PHIL. 3, 39–43, 45–49 (2000); Stephen J. Morse, *Uncontrollable Urges and Irrational People*, 88 VA. L. REV. 1025, 1064–75 (2002).

27. MODEL PENAL CODE § 2.09(1).

28. See Stephen J. Morse, *Excusing and the New Excuse Defenses: A Legal and Conceptual Review*, 23 CRIME & JUST.: REV. RES. 329, 349–62 (Michael Tonry ed., 1998).

29. See HERVEY CLECKLY, *THE MASK OF SANITY* 219–21 (5th ed. 1976). Psychopathy must be distinguished from Antisocial Personality Disorder. See Robert D. Hare, *Psychopaths and Their Nature: Implications for the Mental Health and Criminal Justice Systems*, in *PSYCHOPATHY: ANTISOCIAL, CRIMINAL, AND VIOLENT BEHAVIOR* 188, 189–92 (Theodore Millon et al. eds., 1998). The latter is a disorder officially recognized by the American Psychiatric Association and is defined largely by persistent, serious antisocial behavior. AM. PSYCHIATRIC ASS’N, *DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDERS, FOURTH EDITION, TEXT REVISION [DSM-IV-TR]* 701–06 (2000) [hereinafter *DSM-IV-TR*]. Psychopathy is not so recognized, but there are good data to support the validity of the construct, especially among males, and it is routinely used by clinicians. It is defined primarily in terms of the underlying psychopathology, such as lack of the capacity for empathy, that may predispose psychopaths to engage in antisocial behavior. The overlap between the two categories is substantial but not perfect. Many people who engage in persistent, serious antisocial behavior are not psychopaths, and not all psychopaths engage in the type of antisocial behavior that would support the *DSM-IV-TR* diagnosis of Antisocial Personality Disorder.

nal behavior, human society would be vastly more dangerous than it is.<sup>30</sup> The psychopath cannot be guided, however, by the reasons that empathy and conscience provide. Psychopaths can be guided by purely self-regarding prudential reasons, such as the avoidance of pain, but they are morally irrational.<sup>31</sup> The law now treats psychopaths as responsible,<sup>32</sup> but there is a plausible normative argument that psychopaths are “morally insane” and not sufficiently guidable by reason to be held morally responsible. The point here is not to resolve this issue, but to indicate that the guidance approach offers the most sensible framework for resolving such normative debates about the excuses.

The law’s action-guiding function also clarifies the relationship between wrongdoing and harming.<sup>33</sup> No acceptable moral or political theory holds that people have an absolute right to be free from harm. Harmful accidents will inevitably occur despite the exercise of reasonable care. Because only our intentional actions are fully up to us, all we can fairly ask of each other is that none of us should intentionally place fellow citizens unreasonably at risk of harm. Placing others unreasonably at risk constitutes, *at that moment*, a breach of the duties one owes to the person or class of persons that might be harmed by the agent’s action under the circumstances. The degree of breach is dependent on the mental state with which the agent acted and on the amount of harm risked. For example, intentionally creating the risk of harm with the intention that it occur is more culpable than intentionally creating a risk of harm but without the intent that it should occur.<sup>34</sup> No victim is necessarily harmed by a breach, but the potential for the harm the breach threatens is complete at the moment the agent acts.<sup>35</sup> There is surely the possibility

---

30. This is a familiar observation. See, e.g., THOMAS HOBBS, *LEVIATHAN* 377 (C. B. Macpherson ed. 1968).

31. See also DAVID GAUTHIER, *MORALS BY AGREEMENT* 327–28 (1986) (arguing that moral conduct requires agents to have an “affective capacity for morality,” the capacity to have their “emotions and feelings engaged by what they recognize as moral considerations”); cf. SUSAN WOLF, *FREEDOM WITHIN REASON* 121 (1990) (illustrating a psychopath’s lack of moral reasoning).

32. See, e.g., MODEL PENAL CODE § 4.01(2). The Model Penal Code’s provision exempts from the insanity defense people whose disorder is marked by repetitive criminal or antisocial behavior. This would appear to exempt those with Antisocial Personality Disorder, but potentially to permit the insanity defense for psychopaths. Nonetheless, the provision has been understood and interpreted to exclude psychopaths.

33. See FEINBERG, *supra* note 14, at 105–06 (describing the relation between harming and wronging). Failure to recognize this distinction is a powerful source of error and confusion.

34. Cf. Larry Alexander, *Insufficient Concern: A Unified Conception of Criminal Culpability*, 88 CAL. L. REV. 931, 931, 949 (2000) (arguing that purpose and knowledge can be collapsed into recklessness, and that negligence cannot be distinguished from strict liability and should not be considered a culpable mens rea). But see Joshua Dressler, *Does One Mens Rea Fit All?: Thoughts on Alexander’s Unified Conception of Criminal Culpability*, 88 CAL. L. REV. 955, 956 (2000) (challenging Alexander’s thesis).

35. From the “God’s eye” perspective, of course, there are only certainties: an event was going to happen or it was not. The real ex ante probability is thus zero or one. For human beings acting on Earth without perfect foresight, however, future events must always be considered probabilistically. See Larry Alexander, *Crime and Culpability*, 5 J. CONTEMP. LEGAL ISSUES 1, 5 (1994); Heidi M. Hurd & Michael S. Moore, *Negligence in the Air*, 3 THEORETICAL INQUIRIES L. 333, 358 (2002). Some

for *harm*, but the amount of harm ultimately caused adds nothing to the wrongfulness of the agent's conduct—the violation of the duty owed to potential victims.<sup>36</sup> Potential victims and society are clearly wronged by breaches of criminal law prohibitions, even if no harm ensues, as the criminalization of inchoate wrongful conduct, such as attempt and solicitation, and the criminalization of risk creating conduct *per se*<sup>37</sup> indicate. In contrast, even the greatest harms are not wrongs unless the agent breached a duty by creating an unreasonable risk. Thus, the potential victim or class of victims is fully wronged at the moment of the breach. Purely accidental death is the clearest example of a wrongless harm.

The action-guiding function also gives a firm moral basis to the criminal law's concurrence requirement and to the related view that the criminal law should punish for actions and not for character. Criminal law is concerned with potential culpable harmdoing. It does not punish for thoughts because thoughts alone cause no harm, no matter how evil those thoughts may be, and because thoughts are often unintentional. Nor does the law punish for unintentional bodily movements that cause harm because such movements are not actions and cannot be guided by reason.<sup>38</sup> Finally, the criminal law does not punish intentional actions that cause harm unless the actions are culpably performed, because harmdoers do not deserve blame and punishment unless their actions violated a criminal law prohibition. Criminal law requires concurrence because it aims to guide only potentially culpable actions, which are precisely those in which intentional bodily movements are accompanied by culpable mental states.

Even assuming that one could provide a precise definition of character, it is clear that one's character *per se* does not cause harm and that much of one's character is beyond rational control. Only actions cause harm and only actions are potentially fully guidable by reason. Indeed, action must be conceptually prior to character. Actions can be judged morally without knowing anything about the agent's character; character can only be judged morally in light of the agent's actions.<sup>39</sup> The criminal law can guide good action, but it is mostly powerless to guide good character and it rightly does not punish for character alone. Moreover, what-

---

events at the microphysical level may be indeterministic and essentially probabilistic, but we can safely leave this specter aside for criminal law purposes.

36. See Lawrence C. Becker, *Criminal Attempt and the Theory of the Law of Crimes*, 3 PHIL. & PUB. AFF. 262, 273 (1974) (suggesting that the social harm of crime is "consequent to the *process* of doing the major sorts of conduct we punish criminally").

37. See MODEL PENAL CODE § 211.2 (Recklessly Endangering Another Person).

38. Agents can sometimes be responsible for intentionally or foreseeably placing themselves in situations in which an unguidable movement of the body might cause harm. See *People v. Decina*, 138 N.E.2d 799, 803 (N.Y. 1956). But many such movements cannot be foreseen or prevented by guidable action, no matter how much one widens the time frame for finding culpable conduct.

39. See George Sher, *Ethics, Character, and Action*, in VIRTUE AND VICE 1, 4–5 (Ellen Frankel Paul et al. eds., 1998); see also Michael S. Moore, *Choice, Character and Excuse*, 7 SOC. PHIL. & POL'Y 29, 29 (1990) (comparing the choice and character theories of excuse and arguing that the former is superior).



ever action an agent performs is in a real sense “in character” for the agent.<sup>40</sup> After all, the agent did it, and presumably others with apparently similar characters placed in similar circumstances would not do it. Even if the action was statistically unlikely for the agent, was not the type of thing this type of agent seems predisposed to do by her character, or the agent was subject to unusually stressful or tempting circumstances, it is still the case that every agent is capable of statistically unlikely behavior that she is not usually disposed to do and to which not all people subject to unusual stresses or temptations respond by offending. The criminal law fairly expects all rational agents to act properly even in the face of unusual circumstances for which the agent bears no responsibility. To punish for character would be unjust; to fail to punish because a wrongdoer otherwise has good character would be equally unjust.

In conclusion, moral and legal rules are action-guiding, at least in large part, and thus they are concerned with reasons for action. Moral condemnation and desert and punishment should, therefore, apply only to phenomena capable of being guided by reason.

### III. RESULTS AND LUCK

This part addresses the role that “result luck”—luck concerning whether intended or foreseen results actually do occur—should play in ascriptions of desert and the imposition of punishment. As part II suggested, only action can be guided fully; results cannot be. Using completed attempts and risk creation as examples, this part argues that the action-guiding account implies that the criminal law should try as much as possible to wring luck out of decisions about blame and punishment.<sup>41</sup>

---

40. This conclusion is memorably invoked in the opening of Nick Hornby's *How To Be Good*. The narrator, a well-intentioned family doctor who works in a London clinic, is speaking:

I am in a car park in Leeds when I tell my husband I don't want to be married to him anymore. . . . I really didn't think that I was the kind of person to say so in a car park, on a mobile phone. That particular self-assessment will now have to be revised, clearly. I can describe myself as the kind of person who doesn't forget names, for example, because I have remembered names thousands of times and forgotten them only once or twice. But for the majority of people, marriage-ending conversations happen only once, if at all. If you choose to conduct yours on a mobile phone, in a Leeds car park, then you can't really claim that it is unrepresentative, in the same way that Lee Harvey Oswald couldn't really claim that shooting presidents wasn't like him at all. Sometimes we have to be judged by our one-offs.

NICK HORNBY, *HOW TO BE GOOD* 1-2 (2001).

41. My account of this question is heavily indebted to the following four contributions, all of which argue that results should not matter: Alexander, *supra* note 35, at 3; Joel Feinberg, *Equal Punishment for Failed Attempts: Some Bad but Instructive Arguments Against It*, 37 ARIZ. L. REV. 117, 119 (1995); Sanford H. Kadish, *Foreword: The Criminal Law and the Luck of the Draw*, 84 J. CRIM. L. & CRIMINOLOGY 679, 680 (1994); Stephen Schulhofer, *Harm and Punishment: A Critique of Emphasis on the Results of Conduct in the Criminal Law*, 122 U. PA. L. REV. 1497, 1600-03 (1974). Joel Feinberg writes that “[e]very bona fide philosopher of law tries his hand at least once at the ancient problem of punishing failed attempts.” Feinberg, *supra* at 117. This part is my attempt and provides a different justification from those traditionally offered.

The next part considers the full doctrinal applications of the action-guiding view that results should not matter to desert. Part V addresses the leading contrary accounts of why results should matter.

Attempt liability and risk creation laws present puzzles that have consistently vexed criminal law scholars. Consider, first, whether there should be a distinction between the desert and deserved punishment of an agent who completes a crime and one who simply attempts the same crime. For example, two agents without justification shoot intentionally at point blank range at the heart of a victim with the intent to kill. One victim is killed; the other is saved because he is unforeseeably wearing a medallion under his clothes that stops the bullet. What justifiable difference in culpability and desert could there be? A more general means to approach the question is through what I term the "videotape thought experiment." Imagine that you have an extremely high speed, high resolution tape of a criminal event. Stop the tape at the last frame before the result occurs. In my hypothetical, the tape would be stopped after the gun is fired and just before the bullet makes contact with the victim's body. Assess the agents' culpability at that moment. Do they seem differentially blameworthy (or dangerous)? Suppose you never knew the outcome. Would you think that their actions indicate differential desert? Now run the tape one frame further, to show the result. Neither agent has done anything further. In one case the victim dies and in the other he lives, but does it really seem that the shooter's desert has changed?

Although the law increasingly treats attempts and completed crimes the same, many jurisdictions continue to draw punishment distinctions, and even those that adopt equivalence generally tend to distinguish the two at the highest level of criminality, such as homicide.<sup>42</sup> It is also true that the attitudes of ordinary people support a distinction,<sup>43</sup> but this is a sociological observation rather than a moral or legal argument.<sup>44</sup> Why should there be any difference at all?

A second and perhaps stranger puzzle concerns risk creation. If a defendant creates a substantial and unjustifiable risk of unintended harm and the harm occurs, the defendant is fully culpable. If the harm does not occur, however, even attempt liability does not obtain because attempt doctrine virtually everywhere requires that the defendant intend the harm intentionally risked.<sup>45</sup> At most, the defendant will typically be guilty of some type of much lesser, essentially regulatory crime, such as dangerous driving or "endangerment" crimes generally. For example, imagine two drag-racing agents who drive into facing traffic at an extraordinarily high speed. In one case an approaching driver is killed; in the second, death is narrowly averted by the unforeseeably great skill of

---

42. See MODEL PENAL CODE § 5.05(1).

43. PAUL H. ROBINSON & JOHN M. DARLEY, *JUSTICE, LIABILITY, AND BLAME* 13–28 (1995) (providing an empirical investigation of ordinary people's responses to different doctrines of attempt and completed crime).

44. Prima facie, I take our current practices very seriously indeed, as part IV indicates. But there should be convincing reasons to support those practices, or they should be reformed or abandoned.

45. See JOSHUA DRESSLER, *UNDERSTANDING CRIMINAL LAW* 384–86 (3d ed. 2001).

the approaching driver who avoids the crash by an exceptional maneuver. What justifiable difference in culpability could there be? The videotape thought experiment may also be applied here. Stop the tape just before the impact or the exceptional maneuver. Is the second drag racer less blameworthy or dangerous? In the first case, the defendant will be guilty of involuntary manslaughter—or even murder—and a heavy sentence of imprisonment will be imposed; in the second, the defendant will be guilty only of dangerous driving, or negligent or reckless endangerment, and will do little if any time in prison. Once again, ordinary sentiment supports the distinction, but why should this matter?<sup>46</sup> Why shouldn't the intentional creation of substantial and unjustifiable but unrealized risk be punished the same as realized risk under the same conditions?

The standard argument for why results should not matter is that they are not under the defendant's control and are just a matter of so-called moral luck, which is not a rational foundation for desert.<sup>47</sup> Although I am in sympathy with this argument, more needs to be said about luck and the alleged distinction between action and results. To begin, "luck" is of course a slippery concept. In a causally regular universe, and viewed simply from the standpoint of causation, luck does not exist. All events can be lawfully explained by the sufficient conditions that cause them.<sup>48</sup> On the other hand, if one draws the distinction between events within our conscious control and those that are, practically speaking, matters of chance—both of which are fully caused—we do refer to the latter as matters of luck.<sup>49</sup>

As many have noted, if luck is inconsistent with desert, it is a problem that applies far more broadly than to results alone.<sup>50</sup> The kind of person you are and the opportunities you have are largely, if not entirely, a matter of luck, and seem to be as little up to you as the results of your intentional actions. If you are not predisposed to wicked conduct or if opportunities to express your wicked predisposition seldom arise, you are not likely to offend, but if you unluckily are so predisposed and are unluckily confronted with temptation to satisfy your predisposition, you are highly likely to offend. One can try intentionally to modify one's

---

46. ROBINSON & DARLEY, *supra* note 43, at 28–33.

47. Kadish, *supra* note 41, at 689–90, 699–702 (suggesting, however, that perhaps the emphasis on results is too ingrained in our total moral outlook to permit abandoning it in criminal law); Kimberly D. Kessler, *The Role of Luck in the Criminal Law*, 142 U. PA. L. REV. 2183, 2184–95 (1994).

48. See *supra* note 35. Again, we can safely ignore possible microphysical indeterminacy.

49. See RONALD DWORKIN, SOVEREIGN VIRTUE: THE THEORY AND PRACTICE OF EQUALITY 285–303 (2000) (insisting on the importance of distinguishing the results of chance, for which we are not responsible, from the results of choice, for which we are responsible, because the distinction is crucial to our conception of morality and of ourselves).

50. See, e.g., MICHAEL MOORE, PLACING BLAME: A THEORY OF CRIMINAL LAW 233–46 (1997); THOMAS NAGEL, MORTAL QUESTIONS 24–38 (1979); David Lewis, *The Punishment That Leaves Something to Chance*, 18 PHIL. & PUB. AFF. 53, 55–56 (1989); Galen Strawson, *Luck Swallows Everything: Can Our Sense of Free Will Be True*, TIMES LITERARY SUPPLEMENT, June 26, 1998, at 8. I discuss this issue further below. See *infra* Part VII.

character or to arrange one's life to avoid illicit opportunities, but both can be notoriously hard to do, especially if you are unluckily the sort of person who is not insightful, reflective, effective, and the like.<sup>51</sup>

It should be clear, however, that these so-called problems of moral luck are simply re-descriptions of the metaphysical problem that determinism or mechanism allegedly raise for responsibility generally. In a deterministic or universally caused world of mechanism, the causal antecedents to actions, actions themselves, and the results of actions are all sufficiently caused by the causal laws of the universe operating on antecedent events. Moreover, even if the universe is not entirely caused or determined, it is massively regular—it operates according to what Galen Strawson terms the “realism constraint.”<sup>52</sup> Any sensible theorizing about causation and responsibility must accept this constraint. As a product of causation, then, there is nothing distinctive about the luck of results. For perfectly explicable causal reasons, the causal forces of the universe either do or do not produce the harms a malefactor intends or risks.<sup>53</sup> And for similarly explicable causal reasons, we have the types of characters that we do, we have the opportunities that we do, and we intentionally act as we do.

If all events are caused by and are thus products of “luck,” then such luck applies to everything, including character, opportunity, actions, and results. Consequently, the concept of moral luck may seem to imply that any criterion for desert will be a matter of luck and moral and legal responsibility will be obliterated. Many take this position as a matter of metaphysical truth.<sup>54</sup> For the criminal law, however, causation is not and could not be the issue. A societal institution for guidance and control that was unmoored from concepts of fault and responsibility would not be a system of criminal blame and punishment. It would simply be a consequential regime for controlling dangerous agents. The question is whether a criterion of desert appears morally arbitrary from within our practices of responsibility and punishment.

In brief, there are three main types of responses to the metaphysical challenge to responsibility that determinism or mechanism is thought to

---

51. I have previously given an account of many of the characteristics that help agents “fly straight.” See Stephen J. Morse, *Culpability and Control*, 142 U. PA. L. REV. 1587, 1605–10 (1994). See generally Peter Arenella, *Convicting the Morally Blameless: Reassessing the Relationship Between Legal and Moral Accountability*, 39 UCLA L. REV. 1511, 1609–17 (1992) (arguing, inter alia, that moral responsibility includes the capacity to manage those aspects of one's character that impair the ability to act properly).

52. Galen Strawson, *Consciousness, Free Will, and the Unimportance of Determinism*, 32 INQUIRY 3, 12 (1989).

53. As a product of the lawful regularity of our universe, at least at the macro level, the objective likelihood that any event will occur is always zero or one. See sources cited *supra* note 35.

54. See DERK PEREBOOM, *LIVING WITHOUT FREE WILL* 127–57 (2001). See generally Galen Strawson, *The Impossibility of Moral Responsibility*, 75 PHIL. STUD. 5, 12–16 (1994) [hereinafter Strawson, *Impossibility of Moral Responsibility*].

present.<sup>55</sup> The first, “hard determinism,” claims that responsibility and determinism are incompatible and that determinism is true. Therefore, genuine or ultimate responsibility is impossible.<sup>56</sup> Hard determinism cannot explain retrospectively evaluative criminal law responsibility theories and practices or most moral responsibility theories.<sup>57</sup> It provides an external critique of criminal law that obliterates responsibility and deeply held principles of fairness.

The second response, “metaphysical libertarianism,” agrees that responsibility and determinism are incompatible, but it also claims that human beings—or, at least normal adults—are not determined. Persons allegedly have a capacity for a freedom that permits us to act unencumbered by the causal processes of the universe.<sup>58</sup> For the metaphysical libertarian, the buck stops with us. Libertarianism is mostly consistent with the criminal law’s responsibility practices and doctrines. After all, if the causal influences of endowment luck, character luck, and all the other preact influences can be overridden by contra-causally free action, then there is clearly a distinction between responsibility for action and responsibility for the luck that precedes and follows one’s action. The cost of adopting this apparently elegant solution is that one must adopt a panicky and exceptionally implausible metaphysics in a material universe.<sup>59</sup> Quite simply, it is too metaphysically insecure to ground blame and punishment.

The third response, “compatibilism” or “soft determinism,” is willing to concede that determinism is probably true, but holds that responsibility is possible in a determined universe.<sup>60</sup> Compatibilists correctly claim that adult human beings possess the type of general capacities generally thought to ground ordinary responsibility, such as the capacity to grasp and be guided by reason.<sup>61</sup> They also claim that determinism (or indeterminism) does not explain either responsibility or the excuses. Many compatibilists also believe that morality is a human construction that cannot be justified by appeal to an external, mind-independent

---

55. The full explanation is found in the appendix. In this section of the article, I simply give sufficient detail to motivate the following arguments.

56. See, e.g., PEREBOOM, *supra* note 54, at 127–57; RICHARDS, *supra* note 20, at 135–47; SAUL SMILANSKY, *FREE WILL AND ILLUSION* 40–73 (2000). See generally Strawson, *Impossibility of Moral Responsibility*, *supra* note 54, at 5.

57. See, e.g., R. JAY WALLACE, *RESPONSIBILITY AND THE MORAL SENTIMENTS* 54–61 (1994).

58. See, e.g., ROBERT KANE, *THE SIGNIFICANCE OF FREE WILL* 3–22 (2002). This type of freedom is sometimes called “contra-causal freedom,” “agent origination,” and other terms such as “prime mover unmoved,” meant to convey the flavor of this godlike power.

59. See, e.g., BOK, *supra* note 5, at 1–51 (1998) (arguing that libertarianism is conceptually incoherent); PEREBOOM, *supra* note 54, at 1–88 (arguing that libertarianism is conceptually coherent, but scientifically implausible); Peter Strawson, *Freedom and Resentment*, in *FREE WILL* 59, 80 (Gary Watson ed., 1982) (using the term “panicky”).

60. See, e.g., BOK, *supra* note 5, at 6–29; WALLACE, *supra* note 57, at 58–62; sources cited *infra* note 207.

61. See DANIEL DENNETT, *FREEDOM EVOLVES* 9–13 (2003) (providing a naturalized, evolutionary account of these capacities without using the term compatibilism).

source of moral authority.<sup>62</sup> Compatibilism, which is probably the dominant response among philosophers, thus furnishes the most metaphysically plausible internal justification of responsibility in law and morals.

The criminal law operates within the realm of practical reason. Within that realm, only the compatibilist view provides a potentially satisfactory answer to how action can be distinguished from endowment, opportunity, and results; and thus makes sense of the intuition that luck—commonly understood as determined events that are morally arbitrary—should not matter to desert. From the vantage point of practical reason, the compatibilist, action-guiding account of responsibility does not require god-like, contra-causal freedom for genuine responsibility; and it does not deny the thoroughgoing causality that a material conception of the universe appears to demand. Instead, it holds that the general capacity for rationality is the criterion for responsibility and desert, and it assumes, consistent with ordinary observation and common sense, that all minimally rational adults retain the general capacity and the underlying ability to be guided by good reason, whether or not determinism or mechanism is true. Compatibilism and common sense agree that an agent retains the capacity to be guided by reason even if an agent unluckily has predispositions or traits that lead the agent to ignore or undervalue the dictates of reason, even if the agent unluckily is exposed to the types of situations that make it most difficult to be guided by reason, and even if we concede that character and opportunity are largely and perhaps entirely matters of luck for which the agent is not responsible.<sup>63</sup> In sum, the compatibilist action-guiding account has the elegance of libertarianism without the panicky metaphysics. Unlike hard determinism, the guidance approach comports positively with our practices and normatively with principles of fairness that we endorse.

Results should not matter to desert, because good reason in general and legal rules in particular can fully and directly influence only intentional action. Results are properly objects of celebration and regret, but only actions should be objects of moral praise and blame. Of course, moral and legal rules ultimately aim to prevent unjustifiable harms. Risking harm is wrong precisely because harm is risked. But there is no further wrong, no further violation once the agent acts wrongfully to put a victim at risk. The guiding reasons for and against an action are clearly the reasons for and against successfully achieving the action's goal, because agents typically act with the purpose of achieving some goal by the action. But those reasons can guide only the action and not the ultimate outcome.

We can use reason and our knowledge about how the world works to guide our intentional actions to produce desired results, but we cannot depend on the world to cooperate. Our understanding that the world is

---

62. See, e.g., WALLACE, *supra* note 57, at 87–95.

63. See *supra* Part II.

massively lawful and regular and our knowledge of the laws and regularities inform us what results we can reasonably expect to follow from our actions. But agents do not know the future when they act and cannot know if some intervention will prevent success, even in cases when success subjectively appears certain.<sup>64</sup> Because future success is never assured, virtually all crimes are therefore essentially crimes of risk creation.<sup>65</sup> Events and other agents frequently intervene to prevent the occurrence of results that are entirely expectable *ex ante*. Once we have set in motion through our intentional actions the causal processes that are expected to produce results, we can do no more. We have done all that we possibly could on that occasion to cause the result.<sup>66</sup> At most, we know that our actions can guarantee only the possibility of a result occurring and not the result itself. Anything beyond our action is not fully up to us, is no longer subject to guidance by reason, and therefore cannot rationally be the basis for desert. Again, only actions are potentially fully guidable by reason, and they may be evaluated for desert based on the state of affairs the agent was trying to achieve by the action.

The implicit normative assumption of the foregoing discussion is that we should try to wring luck out of criminal liability as much as possible. But isn't this assumption inconsistent with the great role we allow luck to play in the allocation of benefits and burdens in our law and society in general? Suppose that a scientist using flawed theories and methods nonetheless unintentionally and luckily discovers something really important that took no skill or expertise whatsoever to identify once it was right under her nose. We would all agree that the scientist was just "lucky," but she still collects whatever benefits flow from the discovery, including the reputational reward of being the discoverer. Or, most obviously, consider state lotteries. What these examples and limitless others show is that indeed we do allow luck to play a large role.

It is a commonplace that we think that people deserve to be rewarded for luckily causing good results, even if the good result is the product of "pure" luck or of flawed reasoning and methods. For various reasons concerning incentives and the like, we may wish to reward lucky good results, but I believe that such rewards are not genuinely deserved unless the good results were intended in the appropriate way. Consider a person who engages in terribly endangering behavior that by good fortune produces an entirely unintended and unforeseen benefit. For example, suppose an embittered rival of a terrorist kills the terrorist just as,

---

64. Success may be guaranteed in cases of nonresult crimes in which a crime may potentially be completed simply by the movement of one's body coupled with an appropriate *mens rea*. In such cases, trying and succeeding thus merge.

65. See Alexander, *supra* note 34, at 931-32.

66. If we observe that our action is not well-designed or well-executed to produce the result, or that some intervention will prevent it from causing the desired result, then in some instances we can act intentionally again to produce the result. But, as always, the world may not cooperate, no matter how rational and informed we are and no matter how hard we try.

completely unbeknownst to the rival, the terrorist was about to engage in a homicidal terrorist act.<sup>67</sup> We should all be grateful of course and celebrate the outcome, but does the lucky malefactor deserve our thanks? I think not, and indeed, believe the rival deserves moral condemnation.

Although we understand that many of life's outcomes are irreducibly subject to variables not guided by reason and are not proper indicators of desert, for many consequential reasons we have accepted lottery-like outcomes. Luck is allowed to play an important role morally and legally in many contexts, but I believe that criminal blame and punishment are different and should be distinguished. They are the most stigmatizing and afflictive impositions of state power, as the rules of criminal procedure—such as the higher burden of persuasion—recognize. Whether or to what degree people should be branded as criminals and imprisoned, or even put to death, should not be a matter of a lottery. People should never be punished more than they deserve, even if we allow other benefits and burdens to obtain based on factors other than desert.

Perhaps there is no need to punish equivalently all those who attempt to cause prohibited results because all who try have placed themselves in a punishment lottery.<sup>68</sup> Optimum deterrence does not require that we punish everyone maximally, of course, so we should save the most grievous inflictions for those who cause the feared results. Although the desert of those who succeed may be indistinguishable from the desert of luckier agents who try but fail, the former may be singled out because they entered the lottery and ran the risk of bad luck. The proposed penal lottery is not, of course, a genuine random lottery, which would require all who try to cause a result simply to draw straws every time the result occurs.<sup>69</sup> Some who actually caused the result would then receive the lesser punishment; some who did not would receive the greater. This would strike many people as absurd and unfair, but the pure lottery is not morally distinguishable from the proposed scheme. In addition, the lottery is a blatant denial of equal treatment that does not seem to be consequentially justifiable. Penal desert should not be a matter of either good or bad luck. As Richard Parker observes, "Fortune may make us healthy, wealthy, or wise, but it ought not determine whether we go to prison."<sup>70</sup>

---

67. I discuss subjective and objective views of justification more fully *infra* Part IV.F.

68. *But see* R. A. Duff, *Auctions, Lotteries, and the Punishment of Attempts*, 9 *LAW & PHIL.* 1, 17–30 (1990); Lewis, *supra* note 50, at 59–62.

69. *See* David B. Hershenov, *Punishing Failed Attempts Less Severely than Successes*, 34 *J. VALUE INQUIRY* 479, 484 (2000) (rejecting the fairness of the lottery system).

70. Richard Parker, *Blame, Punishment, and the Role of Result*, 21 *AM. PHIL. Q.* 269, 273 (1984).



## IV. DOCTRINAL APPLICATIONS

In tort cases, results and causation are crucial to decisions about who should bear the loss.<sup>71</sup> Criminal liability, however, is not a predicate for damage awards, and the criminal law need not give results or *ex post* understandings independent substantive significance. This part considers doctrinal applications of the action-guiding account of criminal desert and the claim that results should play no role in ascribing desert. An implication of this conclusion is that perhaps results play an important role because the law implicitly overestimates the value of result evidence for evaluating desert, and not because explicitly good arguments demonstrate that results do contribute independently to desert. I do not aim for comprehensive treatment of each doctrine explored, but instead consider these doctrines as examples of the implications of the action-guidance account.

The first section argues that the law should not distinguish the culpability of agents who commit last-act attempts from those who complete their crimes. Preparatory attempts are often criminalized, but they manifest less culpability and should be punished less than last-act attempts and completed crimes. The following section considers whether the creation of sufficiently great risk of prohibited harm in situations when the risked harm does not occur is as culpable as the creation of the same risk that is realized. I conclude that there should be no difference. The next section argues that because results do not contribute independently to desert, the criminal law can fairly abandon causation of results as an element of crime. The following section suggests that the action-guiding approach explains why accomplices should be held liable, but also why current American punishment practices are not fair. The next section addresses strict liability doctrines that depend on unintended consequences; the section claims that such doctrines are unfair because they cannot properly guide action. The penultimate section argues that in cases involving potential justifications, the use of God's eye, or *ex post* understandings of the circumstances that actually obtained, is an unfair means to ascribe blameworthiness because the agent should be judged only in light of his or her *ex ante* behavior. The last section addresses briefly and generally the epistemic value of results or *ex post* understandings for proving culpability, a question of practical importance deferred in the preceding sections. I conclude that results may sometimes have probative significance, but less so than might be assumed, and that the general lack of probative value of results evidence strengthens the argument that results do not have independent status as moral contributors to desert.

---

71. Results might not matter to liability in the civil system if a society adopted a pure social insurance scheme to compensate victims of noncriminal acts that produce harms.

### A. *Attempt Liability*

This section proposes that last-act attempts should be blamed and punished the same as completed crimes and that both last-act attempts and completions should be distinguished from preparatory attempts, which should be blamed and punished less harshly. The section then suggests that the subjectivist, action-guiding account provides the best understanding of other attempt issues, such as impossible attempts, inherent impossibility, and renunciation or abandonment.

#### 1. *Last-Act, Completion, and Preparation*

The reason results should not matter may seem simple, but it creates a doctrinal difficulty. To satisfy the act doctrine for attempt liability, conduct must go beyond “mere preparation,” but it need not be the last act necessary to cause the prohibited harm. A substantial step or dangerously proximate behavior will be sufficient. Despite the vagueness of such tests, they do have attractive features. For example, sufficiently threatening behavior may allow law enforcement to intervene early to prevent prohibited harms, may under some circumstances provide part of the justification for preventive civil commitment, and often may manifest the firmness of the agent’s criminal purpose. Nonetheless, such doctrines threaten to smuggle too much luck back into desert. Until the last act is performed, events beyond the agent’s control, including the agent’s own thoughts and feelings, can intervene to avert a course of conduct that would apparently inexorably lead to the last act. Some interventions, such as unexpected guilt feelings, may appear to cast credit on the agent; others, such as the appearance of police officers, may not. In either case, however, it is just a bit of luck, something that is not guided by reason, that occurs. It is never certain that preparatory behavior, no matter how substantial, will lead to the last act. Therefore, an agent deserves full blame and punishment, without regard to resulting harms, only in those cases when the last act has been performed.

The last act occurs when the agent has done an act sufficient to cause the harm in the course of ordinary events if the surrounding circumstances were as the agent believes them to be.<sup>72</sup> Performing the last act but failing to cause the prohibited harm should be treated as a completed crime and not as an attempt.<sup>73</sup> Finally, note that it does not follow that the punishment for last-act attempts should be raised to that imposed for completed crimes. The analysis is perfectly consistent with

---

72. The last qualification to the definition, concerning the surrounding circumstances, is necessary to accommodate cases of impossible attempts and inherent impossibility. See *infra* Part IV.A.2–3.

73. See R. A. DUFF, CRIMINAL ATTEMPTS 119–20 (1996); see also Lawrence Crocker, *Justice in Criminal Liability: Decriminalizing Harmless Attempts*, 53 OHIO ST. L.J. 1057, 1066 (1992).

lowering the punishment for completed crimes to the level now imposed on attempts.<sup>74</sup>

One might claim that some last act attempters are less culpable because they know that they lack the skill to create much risk that they will succeed. For example, consider a person with poor eyesight and shaky hands who tries to shoot his victim from some distance. Holding all else constant, the risk that the victim will be hit and die is considerably less than if the shooter were an expert marksperson. In some cases at the extremes, there might be some merit to this argument. At very low levels of skill, and if the defendant is aware of how much skill is lacking, intending a result is more like consciously creating a risk of that result and hoping that it will be realized. Perhaps these cases should be analyzed as risk creation cases. The question would then be whether sufficient risk was created by the defendant's plainly reckless actions.<sup>75</sup> If the defendant's skill is not so low, however, or if the defendant does not know that his or her skill is so low, then the defendant may fairly be held fully liable for the last act attempt to commit the crime, even if more skillful agents would be more certain of success. Lack of skill is not a general defense to attempt liability.

One might also claim that those last act attempters who fail are less culpable and perhaps less in need of incapacitation or reform because failure indicates that the agent was not wholeheartedly committed to the criminal goal.<sup>76</sup> As a positive matter, however, current criminal law grades neither crimes nor punishments according to the agent's wholeheartedness. Even if we thought as a normative matter that it should do so, there would be no reason to use success or failure as anything other than an evidentiary consideration bearing on wholeheartedness. After all, a criminal with reservations might succeed and a criminal without reservations might fail. Moreover, assessing wholeheartedness and motivation generally will be a difficult enterprise, especially if one uses speculative psychodynamic formulations about unconscious processes to do so.<sup>77</sup> Success or failure of a last act attempt is unlikely to be a good proxy for wholeheartedness. It certainly will not be sufficient fairly to ground a substantial differential in punishment.

---

74. In general, I believe that most criminal sentencing schemes in the United States are unfairly and unnecessarily harsh.

75. The guidance account suggests that reckless creation of sufficient risk should be treated the same as the reckless realization of the same risk. See *infra* Part IV.B (discussing risk creation).

76. See Becker, *supra* note 36, at 288; Hershenov, *supra* note 69, at 482–83.

77. For example, a foiled bank robber might try to introduce evidence that he unconsciously wished to fail. See, e.g., *United States v. Pollard*, 171 F. Supp. 474, 480 (E.D. Mich. 1959). In addition to being scientifically speculative, such accounts usually try to smuggle in without argument the allegedly entailed conclusion that the defendant's behavior was uncontrollable or that the defendant lacked "free will." See generally Stephen J. Morse, *Failed Explanations and Criminal Responsibility: Experts and the Unconscious*, 68 VA. L. REV. 971, 1082–83 (1982) (examining the conceptual and scientific difficulties psychodynamic formulations present and concluding that such formulations should not be admissible in criminal trials).

If the intentional last act is sufficient to warrant full blame and punishment without regard to results, then is preparatory behavior—incomplete attempts—therefore beyond the scope of moral and legal condemnation? Are preparatory acts, coupled with the intent to complete the crime, too subject to endowment and opportunity luck to justify punishment for such acts? I confess to ambivalence about this question, but there is a plausible argument for criminalization of incomplete attempts. Although events may intervene to prevent agents from fully performing the last act, as agents come closer to the last act in an intended course of conduct, in general they tend to become increasingly committed to completion and the potential for intervention decreases. Most preparatory conduct is not itself terribly dangerous or immoral, but agents know that they should not step onto the path of evil, because having done so, each next step becomes easier. Such knowledge gives them good reason not to proceed. Thus, the risk that the last act will be performed increases as the agent intentionally performs acts that increasingly approach the last act.

At some point prior to the last act, the agent's intentional conduct may be sufficiently immoral and dangerous per se to warrant condemnation and punishment, because some conduct creates an undeniably substantial risk that the agent will in fact perform the last act. I assume that this point would be "dangerously proximate" to the last act because otherwise luck would play a role more powerful than desert permits. Thus, the law should favor attempt act tests that come close to completion. Such tests, by giving citizens clearer warning, have the further virtue of better satisfying the principle of legality than tests that criminalize conduct more remote from the last act.<sup>78</sup> But agents engaged in preparatory conduct should not be blamed and punished as much as those who perform the last act, because preparatory attempts do not fully risk the prohibited result. There is simply too much opportunity for genuine luck to intervene.

On the other hand, because purpose must always be conditional prior to the last act and preparatory conduct is rarely harmful in itself, perhaps punishment for incomplete attempts is punishment for thoughts rather than for culpable actions. This view also seems plausible to me. Note that if incomplete attempts were no longer criminalized, the only forms of restraint that could be preventively exercised would be independent criminal prohibition of sufficiently dangerous preparatory conduct, such as possession of instruments of crime<sup>79</sup> and civil commitment.<sup>80</sup> Criminalization would not be a strong substitute for preparatory attempt

---

78. See GEORGE P. FLETCHER, *RETHINKING CRIMINAL LAW* 181 (1978).

79. See MODEL PENAL CODE § 5.06(1) (1962).

80. See Stephen J. Morse, *Neither Desert nor Disease*, 5 *LEGAL THEORY* 265, 267–70 (1999) (arguing that restraint is justified only if an agent performs a culpable act or threatens to perform a dangerous act but is not responsible).

liability, however. Much obviously preparatory conduct is not dangerous enough to warrant substantial punishment or even to warrant criminal prohibition at all. Furthermore, because most agents engaged in incomplete attempts do not suffer from a mental abnormality or any other problem that diminishes responsibility, civil commitment would not be a justifiable form of preventive restraint. Thus, there is a good consequential argument for maintaining the criminalization and punishment of incomplete attempts. Because there is also a plausible claim that incomplete attempts are culpable acts, let us assume that the case for desert is sufficient to justify punishment.

The law of attempts now draws the line between attempts and completed crimes based on whether the prohibited harm occurs. My analysis suggests that culpable criminal conduct is complete when the last act is performed, and that results should play no role in desert or punishment. This analysis also suggests that there is wisdom in distinguishing incomplete attempts from both completed crimes and completed attempts if the line between them is drawn properly.<sup>81</sup> If luck is largely expunged from the course of criminal conduct *before* as well as after the agent does the last act, then the preparing agent should be blamed and punished less severely, because preparatory conduct is simply not as risky and dangerous as the last act. Moreover, the agent can then be guided by the punishment differential that provides an incentive to desist from further preparatory conduct.<sup>82</sup>

The conclusion that last acts should be treated as morally and legally indistinguishable from completed crimes potentially raises a final, practical problem. The act element of crimes that do not require results and the result element of result crimes are clearly defined. For example, whether the property was taken and carried away or whether there is a dead body is seldom difficult to decide. But in some cases there may be uncertainty about whether the agent has done the last act, an act sufficient in the ordinary course of events to produce the prohibited harm.<sup>83</sup> Perhaps the criminal justice system should not expend the resources to distinguish last acts from mere preparatory attempts. Whether this will be a costly task is an empirical matter about which there are no data, but

---

81. Drawing the line in this way would also provide a better moral and intuitive justification for allowing a renunciation defense to attempt. Renunciation is justified primarily on consequential grounds instead of desert, however. See *infra* Part IV.A.4 (discussing renunciation).

82. This is the only consequential argument for the punishment differential that seems to have credible purchase. It does not appear to be a substantial consideration, however, and I raise it only for completeness. See generally Schulhofer, *supra* note 41, at 1517-87 (canvassing and rejecting virtually all differential punishment arguments based on deterrence).

83. This question must be distinguished from the question of whether one can accurately infer the agent's mental state from the agent's action, whether the action is preparatory or a successful or unsuccessful last act. The question must also be distinguished from the question of culpability in some cases of inherent impossibility, in which the last act by definition is insufficient in the ordinary course of events to cause the prohibited harm. In such cases, we may wonder if the defendant contemplated doing further conduct that would be sufficient. In virtually all cases, however, it will be clear that the defendant did all that he or she thought would be sufficient.

I suggest that problematic cases will be rare. Moreover, identifying what counts as the last act is substantially less difficult and uncertain than the apparently acceptable uncertainty attending the identification of sufficient acts under the immensely vague present tests that criminalize and punish preparatory behavior. Furthermore, questions about the last act will arise only in result crime cases because last acts and completions collapse together in nonresult crimes. Finally, any costs incurred in developing and applying a last act test will be outweighed by the benefits of abandoning the evaluation of causation in criminal cases.<sup>84</sup>

## 2. *Impossible Attempts*

The action-guiding approach also makes sense of the perennially thorny problem of impossible attempts. In all cases of impossibility, the agent has committed the last act. Failure occurs, however, not because the agent's performance errs or some other variable actively intervenes, but because success was impossible *ex ante* on that occasion. For example, the agent shoots the potential victim with the intent to kill and the bullet pierces the victim's heart, but the victim is already dead.<sup>85</sup> From the vantage point of practical reason, impossible attempts are indistinguishable from other last-act completed attempts because in all cases the agent has tried to produce a prohibited harm and has done what would be a sufficient act if the surrounding circumstances were as the agent believed them to be.<sup>86</sup> Whether failure is produced by poor performance, active intervention, or unknown states of affairs is irrelevant to the law's ability to guide the agent's conduct.<sup>87</sup> In cases of impossible attempts, it may be difficult to determine whether the agent intended to produce the prohibited harm, because the agent's manifest conduct may be innocent. For example, suppose an agent takes and carries away an umbrella believed to belong to another with the intent permanently to deprive, but the umbrella in fact belongs to the agent.<sup>88</sup> In such cases, it might be exceedingly difficult to prove the necessary *mens rea*, but the difficulty is

---

84. Part IV.C argues that causation elements may be abandoned in criminal law. See *infra* Part IV.C.

85. See *People v. Dlugash*, 363 N.E.2d 1155, 1162–63 (N.Y. 1977).

86. We can distinguish those cases the law terms “impossible” from other failed attempts. In the former, in principle we could know *ex ante* if success was impossible; in the latter, in most cases we only can know *ex post*. Note that from the God's eye vantage point, all failed attempts were “impossible” *ex ante*. What did happen is the only event that could have happened.

87. From the vantage point of theoretical reason, the only events possible are the unique events determined by preceding events and the laws of the universe. Metaphysically, then, “impossible” attempts are indistinguishable from those that fail because either the agent performs poorly or some other variable intervenes. In this view, however, all failures and successes are indistinguishable.

88. SANFORD H. KADISH & STEPHEN J. SCHULHOFER, *CRIMINAL LAW AND ITS PROCESSES* 600 (7th ed. 2001). For a more realistic but still problematic case, see *United States v. Oviedo*, 525 F.2d 881, 882 (5th Cir. 1976) (defendant sold an uncontrolled substance that he said was heroin to an undercover agent and was charged with attempted sale of a controlled substance; defendant claimed that he knew the substance was uncontrolled and intended only to “rip-off” the buyer).

purely epistemic. The agent is morally a last-act attempted umbrella thief and the agent's desert is indistinguishable from that of those who succeed.

### 3. *Inherent Impossibility*

The guidance account also provides a sensible approach to cases of so-called inherent impossibility, those in which the agent commits a last-act attempt, but success is impossible because the agent uses means utterly ill-adapted to achieving the prohibited harm. For example, imagine a defendant who tries to crack a bank safe using the beam from an ordinary flashlight. Unlike the case of standard impossibility, success was possible only if one suspends the causal laws of the universe. If the agent is just foolish or careless, then the agent is fully capable of being guided by reason and the law. Consequently, the agent is a fully culpable last-act attempter and should be treated the same as other last-act attempters and as those who successfully complete the crime. In some cases, however, the agent's mistake may indicate that the agent is not a rational agent and therefore cannot be guided. If so, such agents should be excused.

### 4. *Renunciation or Abandonment*

Finally, the action-guiding approach explains why the renunciation or abandonment defense to attempt liability can be justified by a consequential theory of punishment but not by desert.<sup>89</sup> The defense is permitted only if the defendant renounces because he or she has had a purely internal change of heart and was not caused to abandon the criminal conduct by fear of detection or the like.<sup>90</sup> But the action that the agent has already done is no less culpable because the agent later "gets religion." Suppose a miscreant undergoes a genuine moral transformation the moment after a "result" crime, such as homicide, is successfully committed. Assume that there is no question about the agent's sincerity and that henceforth the agent will live a life of sparkling moral purity and selfless sacrifice for the common good. Does the person deserve no blame and punishment on moral grounds? I think the agent is fully blameworthy and deserves punishment because desert judgments are retrospective. One might wish to mitigate or forgo punishment altogether on consequential grounds, to reduce unnecessary incapacitation, for example, but not because the agent was not fully culpable when the crime was committed. Or, as in cases of preparatory attempts, we might wish to provide a defense to give the agent an incentive independently to change his or her mind. But again, the agent who intends to complete is

---

89. See also Morse, *supra* note 80, at 288–89 (discussing renunciation doctrine).

90. See, e.g., MODEL PENAL CODE § 5.01(4) (1962).

fully culpable for the preparatory attempt and would be blameworthy and punishable as soon as the conduct meets the requirement for attempt liability.<sup>91</sup>

Even an agent who successfully completes a crime, but then fully undoes the harm later, does not get the defense. Suppose, while the victim is on vacation, a thief takes and carries away a cherished pink flamingo lawn decoration that graces the victim's front lawn with the intent to deprive the victim permanently of this ravishing object. Later on, after considering the gravity of what has happened, the agent returns the pink flamingo to precisely the spot where it was originally placed. Again, on consequential grounds we might wish to mitigate punishment, but not because the thief was less culpable when he made off with the flamingo. We wish simply to give the agent an incentive to mitigate the harm. The moral evaluation of an action is complete at the time of the action, independently of an agent's later action. We might more favorably evaluate the character of a criminal who later tries to mitigate the harm, but the character of the original criminal act is not changed by later conduct, no matter how admirable it may be.

One might wish to claim that renunciation indicates that the agent's commitment to his criminal conduct was not wholehearted and therefore that the agent was less culpable. As we have already seen, however, even assuming implausibly that we could validly assess wholeheartedness, it is not clear that wholeheartedness affects desert more than trivially, if at all.<sup>92</sup> Behavior that satisfies the elements of a crime is fully criminal, and the victim and society are equally wronged no matter what degree of wholeheartedness existed in the wrongdoer's soul. We might think that agents who wrong others halfheartedly have better characters than agents who act without reservation, but the wrong and culpability for the offenses are equal. In general, character assessment and incentives, and not desert for action, are what drive intuitions in renunciation cases.

### B. Risk Creation

Crimes of risk creation ought also to be complete whenever the agent's intentional action creates the prohibited risk and the agent possesses the appropriate mens rea—recklessness or negligence—concerning the risk. The risk creation cases that fill the appellate reports and criminal law casebooks provide terrifying examples. They tell tales of drag racers, Russian roulette players, parents who brutally abuse or neglect their children, motorists who drive wildly, the creation of ultra-hazardous conditions, and other stories of gross lack of concern for the

---

91. Note that renunciation is possible in last-act attempt cases mostly in the relatively infrequent situations involving "long, slow fuses." The hypothetical cases that Leo Katz uses to try to show that results contribute to desert are of this type. See discussion *infra* Part V.B.

92. See *supra* text accompanying notes 76–77 (discussing wholeheartedness).



safety of fellow citizens that are not for the fainthearted. Whether the risked outcome occurs should not matter in these cases because this fortuity cannot guide the agent's conduct and has no relevance to desert.

The issue is precisely analogous to the case of intended harms. The agent who creates a sufficient risk has at that moment fully breached the legal rule that is meant to guide conduct and deserves to be punished fully. Heidi Hurd has argued that it is viciously circular and deontologically unacceptable to construe creating risks as wrongs,<sup>93</sup> but there is no circularity, even if risking harm is not itself a harm.<sup>94</sup> Harms and wrongs are distinct. Moreover, I see no reason why it is deontologically unacceptable to consider violations of duties to others as wrongs, even in the absence of resulting harms. What could be more obviously wrong in itself than intentionally breaching a moral duty owed to others? In the case of homicide, for example, this analysis suggests that criminal liability for homicidal behavior should be based on pure death-endangerment.

Many might rebel at this conclusion. They might wonder, for example, whether careless drivers should be punished as if they had committed involuntary manslaughter. The answer is that they should, *if* the amount of risk of death they created was sufficient to satisfy the much higher level of risk required to differentiate criminal from tort liability. In turn, I wonder why this answer seems unacceptable. In terms of desert and danger, the sufficiently careless driver who luckily does not kill is indistinguishable from the similarly careless driver who unluckily does. I assume, further, that we wish maximally to deter the kind of gravely death-endangering conduct that is a predicate for vehicular homicide liability, especially if the risk would justify a conviction for murder. This proposal would surely lead to much more careful operation of motor vehicles, and thus to less carnage from automobile accidents—which cause many more deaths than nonvehicular criminal homicides.<sup>95</sup> Among those who drive sufficiently carelessly nonetheless, the vagaries of realistic law enforcement would create a “lottery,” because all careless drivers could not be apprehended by police. Unlike the proposed penal lottery,<sup>96</sup> however, this form of lottery is inevitable and the luck involved has nothing to do with differences between risk-creators; it is entirely a function of police behavior. This lottery would not be unfair as long as the police behave evenhandedly.<sup>97</sup>

---

93. Heidi M. Hurd, *The Deontology of Negligence*, 76 B.U. L. REV. 249, 264–65 (1996).

94. Stephen R. Perry, *Responsibility for Outcomes, Risk, and the Law of Torts*, in PHILOSOPHY AND THE LAW OF TORTS 72, 76–77 (Gerald J. Postema ed., 2001).

95. In 1999, there were 42,437 deaths from motor vehicle accidents and 16,831 deaths from “assaultive” homicide. Kenneth D. Kochanek et al., *Deaths: Preliminary Data for 1999*, 49 DHHS NAT'L VITAL STAT. REP. 3, 5 (2001).

96. See *supra* text accompanying notes 68–70.

97. Police misconduct toward minority motorists does give one pause, however. See David Rudovsky, *Law Enforcement by Stereotypes and Serendipity: Racial Profiling and Stops and Searches Without Cause*, 3 U. PA. J. CONST. L. 296, 299–304 (2001) (reviewing multiple sources of data demonstrating the widespread use of profiling).

Finally, taking per se risk creation seriously would lead to finer calibration of culpability assessments. Treating risk creation as criminal without regard to results would have the virtue of forcing the law to determine the degree to which certain commonly risky activities, such as driving while intoxicated, are in fact sufficiently death-endangering per se to warrant the heavy penalties associated with homicide liability.<sup>98</sup>

### C. Causation

Another doctrinal implication of this account of results is that causation should be irrelevant to criminal liability.<sup>99</sup> After all, causation matters to criminal liability only because results do. I have claimed that results should not matter because criminal law is addressed to action and aims to achieve retributive justice proportional to desert. Unlike tort law, criminal law does not seek to decide where losses should lie for the purpose of compensation. If results do not matter, then we can safely jettison the extraordinarily confusing and vague tests for proximate cause to determine when intervening events “cut the causal chain” and thus exempt the agent from responsibility for the results.<sup>100</sup>

Even on their own terms, causation tests can have nothing to do with desert because they are not addressed to and cannot guide the actions of potential wrongdoers. In all criminal cases raising causation issues, the defendant’s conduct has satisfied the elements of prima facie liability and was a but-for cause of the prohibited result. The defendant has thus done all that he or she could to bring about, or to risk, the prohibited result, and has actually caused the result. *After* the defendant acted, however, something unforeseeably deviant or wacky intervened to complete the causation of the harm. Agents virtually never act with the expectation that an improbable set of events will follow.<sup>101</sup> Indeed, in virtually all cases of intent to cause the harm, the agent will not care precisely how the result was caused, but will be satisfied that ultimately it

---

98. Compare Model Penal Code § 2.08(2) (1962), which would hold an agent guilty of recklessness, even if the agent acted without conscious awareness of the risk created, if voluntary intoxication was the reason the agent was not aware of the risk. The defense of this seeming bit of strict liability—usually anathema to the MPC—is that agents are always actually aware before they become intoxicated of the risks they may create without awareness when intoxicated. But this is utterly implausible. See Stephen J. Morse, *Fear of Danger, Flight from Culpability*, 4 PSYCHOL., PUB. POL’Y, & L. 250, 254 (1998).

99. This section is derived from and builds on an earlier treatment of the same issue. Stephen J. Morse, *The Moral Metaphysics of Causation and Results*, 88 CAL. L. REV. 879, 881–89 (2000).

100. See Hurd & Moore, *supra* note 35, at 336 (referring to proximate cause tests as “elusive, multiple and often conflicting in their implications for cases”). Hurd and Moore think that they can successfully rationalize causation doctrine, but nearly all their results are derived from tort law, not criminal law, and—to be consistent—they defend some of the most indefensible criminal law doctrines, such as the rule that requires criminals to take their victims as they find them. See *infra* Part IV.E (discussing why this doctrine is unfair).

101. If the agent intentionally takes advantage of a statistically unlikely coincidence to achieve his or her end, then the coincidence will not break the causal chain. See H. L. A. HART & TONY HONORÉ, *CAUSATION IN THE LAW* 77–79 (2d ed. 1985).

was caused. After all, the agent acted intentionally to cause that result. Nor will the manner in which a risk is realized *ex post* affect the risk creator's conduct *ex ante*. How the causal events actually unfold is most decidedly not up to the agent and could not possibly affect her behavior.

Causation tests are addressed only to legal decision makers—juries and judges. But on what possible principled ground can they decide whether the agent proximately caused the harm? To the best of our knowledge, causation is a seamless web with no “gaps” or “breaks” in the “chain” that metaphysically correspond to the bright line judgment that proximate cause was lacking. Even the notion of causation “petering out”—the idea that as the causal chain lengthens, earlier causes account for less of the variance—is only a slightly more intuitive but nonetheless inaccurate metaphor. Proximate cause is therefore said to be a policy choice, but in criminal law, what is the policy? Proximate cause seems entirely unrelated to considerations of desert, deterrence, and incapacitation.<sup>102</sup>

Consider, for example, the following traditional causation doctrine. The criminal law is reluctant to trace responsibility for consequences back through the last independent, responsible agent.<sup>103</sup> If the criminal law were not concerned with results, but only actions, however, this potential limitation on liability would be of no importance. Suppose the following case: A perpetrator intentionally or with conscious disregard of an especially grave risk of death inflicts a mortal wound on the victim, who is then rushed by ambulance to the hospital. In one variation, the ambulance driver drives very carefully and reaches the hospital safely, but the victim dies in the emergency room. In this case, the perpetrator is guilty of intentional murder or “depraved heart” reckless murder. In the second variation, the ambulance driver drives wildly and unjustifiably carelessly, the ambulance crashes, and, as a result, the victim's neck is broken, killing him instantly. In this case, the intentional perpetrator is guilty of only attempted homicide; the reckless perpetrator is guilty of reckless endangerment or perhaps some type of assault crime. What possible difference in the perpetrator's culpability could there be between the two scenarios? If results and causation did not matter, then there would be no difference. In criminal law, there is no desert-based principle to provide moral significance to an unintended and unforeseen independent agent who causally contributed to the result. Agents should be judged morally by their own behavior.

If the causal chain seems too deviant, then the ordinary person's rough sense of justice might be offended by attributing the result to the

---

102. I take no position here about the policies supporting proximate cause doctrines in tort. Torts theorists famously disagree about the justification for recovery, but only money is at stake. In criminal law, blame and punishment are crucial, and desert is at least a necessary condition for liability.

103. See HART & HONORÉ, *supra* note 101, at 74.

agent,<sup>104</sup> but once again, this is a sociological observation and not a rational argument. If this rough sense of justice is too ingrained to be abandoned, then we should recognize that any test developed will be rationally arbitrary. If results were jettisoned as a criterion for criminal liability, then the criminal law could avoid the baroque machinations of legislators, lawyers, and judges as they try to rationalize proximate cause doctrine.

In sum, results are properly objects of celebration and regret, but only actions should be objects of moral praise or blame. For compatibilists, this is a fully causal universe. But results should not matter to desert and therefore causation should not either. Our knowledge of how the world causally works should enter into responsibility assessment only by guiding judgments about the type of harm an agent's intentional conduct risked.

#### D. *Accomplice Liability*

Accomplice liability—aiding or encouraging a perpetrator with the intent to promote or facilitate the perpetrator's crime<sup>105</sup>—is a puzzle. The accomplice has demonstrated moral culpability through both a culpable mental state and the action of aiding and encouraging the crime. That much is straightforward. The puzzle is produced by the apparent inconsistency of accomplice liability and the criminal law's more general view of agency and causation.<sup>106</sup> The accomplice does not coerce the perpetrator.<sup>107</sup> Indeed, the perpetrator is fully, independently criminally liable, even if the accomplice is a but-for cause and provides assistance necessary to the perpetrator's success. Conversely, the accomplice's action need not play any causal role whatsoever in promoting or facilitating the perpetrator's deed. The accomplice will be fully liable even if she offered slight assistance or the perpetrator was completely motivated to commit the crime and needed no aid or encouragement. It is sufficient if the accomplice's conduct might, counterfactually, have played some causal role, however trivial it may have been.

Suppose, plausibly, that we treat the perpetrator's criminal behavior as the undesirable result. What is the accomplice's requisite causal role? None. And it is a commonplace of causation analysis that in criminal law we do not trace consequences back through an independent, responsible agent.<sup>108</sup> Why, then, does American law treat the criminal liability of ac-

---

104. ROBINSON & DARLEY, *supra* note 43, at 181–89.

105. See MODEL PENAL CODE § 2.06(3) (1962).

106. SANFORD H. KADISH, *Complicity Cause and Blame: A Study in the Interpretation of Doctrine*, in *BLAME AND PUNISHMENT: ESSAYS IN THE CRIMINAL LAW* 135, 143–46 (1987).

107. If the accomplice sufficiently coerces the perpetrator, the latter becomes an excused “innocent instrumentality,” and the accomplice becomes the perpetrator who acts “through” the innocent instrumentality. MODEL PENAL CODE § 2.06(2)(a).

108. HART & HONORÉ, *supra* note 101, at 74.

complices and perpetrators as equals? Indeed, why are accomplices criminally liable at all?

The usual justification for the accomplice's derivative liability is that the accomplice has identified with or consented to the perpetrator's crime.<sup>109</sup> In essence, the accomplice has made the crime his own by aiding or encouraging conduct. Another justification is "forfeited personal identity:" the accomplice allegedly "forfeits" his or her right to be treated as a separate person.<sup>110</sup> But if the aiding or encouraging conduct is not independently criminal—or if it is much less so than the perpetrator's crime—and if the accomplice need not play any causal role, how can liability be justified by consent or forfeiture? Understanding why accomplices should be held liable at all and how they should be punished requires a justification for accomplice liability different from the usual suspects.

The most promising basis for accomplice liability is simpler and more straightforward than consent, identification, or forfeiture. As a general matter, furnishing aid or encouragement that might play some role increases the risk that the perpetrator will successfully commit the crime. It is simply a fact about human beings that potential encouragement or aid in principle increases the risk that the recipient will commit the act aided or encouraged. This is true even if we treat the perpetrator as a fully independent, responsible agent. Although responsible agents must "take responsibility" for what they ultimately do, other people can play a causal role in enhancing or decreasing the risk of what an agent does. Consequently, even if the potential accomplice is not perpetrating the deed, it is his or her duty not to enhance the probability that the crime will be committed. It is wrongful because it is potentially socially harmful to furnish aid or encouragement with the purpose of facilitating the crime. To accept this argument is not to abandon important moral conceptions of agency; it is simply to recognize how the world works.

The foregoing account of accomplice liability focuses entirely on the accomplice's own behavior and suggests that accomplice liability should not be derivative. This justification for the basis of accomplice liability may also help explain liability for conspiracy and solicitation, which allegedly raise similar problems of agency and causation. The accomplice has done everything that he or she can to aid or encourage the perpetrator. What the perpetrator then does is not within the control of the accomplice. The Model Penal Code and those states that follow it have recognized this point by holding a potential accomplice for attempt even if the potential perpetrator neither commits nor attempts the crime the

---

109. Joshua Dressler, *Reassessing the Theoretical Underpinnings of Accomplice Liability: New Solutions to an Old Problem*, 37 HASTINGS L.J. 91, 110–11 (1985); KADISH, *supra* note 106, at 158–59.

110. Dressler, *supra* note 109, at 111. I find this justification too metaphorical, but I include it for completeness.

potential accomplice was trying to promote or facilitate.<sup>111</sup> I would go further and simply hold the accomplice liable for the crime intentionally facilitated or promoted, subject to the punishment limitations I am about to address. I recognize that at present there are formal limits that would prevent this outcome because liability is considered derivative. There would be no bar to achieving this result by a specific statute, however. For example, the Model Penal Code has gone far beyond the traditional common law by permitting a would-be accomplice to be convicted for attempt.<sup>112</sup>

The account also suggests that accomplice liability should not be automatically equal to the perpetrator's liability. In appropriate cases, culpability should be differentiated according to the individual mens rea of the accomplice and the perpetrator, as it now is for homicide. Otherwise, accomplice liability should be proportionate to the substantiality of the aid or encouragement furnished. In general, the substantiality of the aid will be positively proportionate to the enhancement of the risk that the crime will be perpetrated, and thus to the accomplice's level of wrongdoing and culpability.<sup>113</sup> Few accomplices will likely be as or more culpable than the perpetrator, and those accomplices who know their aid is relatively insubstantial would be considerably less liable even if the trivial accomplice were thoroughly identified with the perpetrator's crime. According to this account, accomplice liability would theoretically be a continuum, but, in practice, a few categories of accomplice liability would be sufficient to do justice in a world of imperfect information.

Let us apply the foregoing suggestions to a famous case that has divided the commentators, *Regina v. Richards*.<sup>114</sup> Isabelle Richards hired two men to beat up her husband, telling them that she "wanted them to beat him up bad enough to put him in the hospital for a month."<sup>115</sup> The men attacked her husband, but he was able to escape without serious injury. Ms. Richards and the two men were charged with wounding with intent to do grievous bodily harm, but the two men were convicted only of the lesser included offense of intentional wounding.<sup>116</sup> As an accomplice, should Ms. Richards be guilty of the greater or lesser offense?<sup>117</sup>

---

111. MODEL PENAL CODE § 5.01(3).

112. *Id.*

113. See FLETCHER, *supra* note 78, at 649-52 (noting that some criminal justice systems always treat accomplices as less culpable than the perpetrator and try to punish them in proportion to their wrongdoing and culpability).

114. [1974] Q.B. 776 (C.A. 1973). For commentary, see *Regina v. Howe* [1987] 1 All E.R. 771, 799 (H.L. 1987); J. C. SMITH, CRIMINAL LAW 172 (10th ed. 2002); KADISH, *supra* note 106, at 181-86; J. C. Smith, *Comment*, CRIM. L. REV. 480, 484 (1987).

115. *Richards*, [1974] Q.B. at 777-78.

116. I do not know why the men were convicted of only the lesser offense. Although it is clear that they intentionally wounded Mr. Richards and intended to do grievous bodily harm, the jury might not have been willing to convict the hirelings for the serious felony in the absence of any serious harm.

117. She is clearly guilty of solicitation and conspiracy.

Once she hired and directed the men, her behavior was complete. What happened thereafter was not up to her. Her mens rea, explicit in the wish that the beating should lead to a month's hospitalization, was clearly for the greater offense. She incontrovertibly intended to enhance the risk that the more serious wounding would be accomplished. How substantial was her aid, however? Although the men were entirely responsible for their own conduct, it would not have occurred if Ms. Richards had not hired and directed them. I would argue that this but-for conduct is very substantial aid and encouragement—indeed, it could hardly be more so to enhance the risk. Ms. Richards should properly be convicted for the higher level of crime. What makes this case complicated is the allegedly derivative nature of accomplice liability. I submit that if accomplice liability were treated as I suggest, focusing on each agent's action as the independent ground for individual desert, the answer is quite straightforward.

#### E. *Strict Liability*

Strict or absolute liability appears in two forms in criminal law: "pure" and "bump-up."<sup>118</sup> Pure strict liability obtains when the definition of the crime contains no fault element, no mens rea. Bump-up strict liability occurs when a defendant culpably engages in criminal conduct that causes consequences other than those the defendant intended or contemplated, and the defendant is held liable for the further consequences automatically and without proof of the usual mens rea required for conviction for the further consequences. Familiar examples of bump-up strict liability include the following: felony murder;<sup>119</sup> the causation rule that the defendant must take the victim as he finds him or her;<sup>120</sup> the accomplice liability rule that holds an accomplice liable for a perpetrator's reasonably foreseeable intentional conduct that the accomplice did not intend to promote or facilitate;<sup>121</sup> and the rule that holds a defendant guilty of a higher offense if the defendant mistakenly thinks that he or she is committing a lesser legal crime but the actual circumstances are sufficient for a higher offense.<sup>122</sup> One might think that the existence of these and similar doctrines undermines my claim that action-guidance is the best positive explanation of desert and the criminal law generally, but

---

118. The labels are my own invention and are used only for convenience.

119. See, e.g., CAL. PENAL CODE § 189 (West 1999 & Supp. 2003); 18 PA. CONS. STAT. ANN. § 2502(b) (West 1998 & Supp. 2003).

120. See, e.g., HART & HONORÉ, *supra* note 101, at 79–80.

121. See, e.g., *People v. Luparello*, 231 Cal. Rptr. 832, 848 (Cal. Ct. App. 1987). An analogous rule applies to coconspirators in those jurisdictions that accept the *Pinkerton* rule. *Pinkerton v. United States*, 328 U.S. 640, 646–47 (1946) (holding that coconspirators may be held liable as accomplices to substantive crimes committed by coconspirators that were not part of the original conspiratorial agreement if those crimes were reasonably foreseeable).

122. See, e.g., *People v. Lopez*, 77 Cal. Rptr. 59, 63 (Cal. Ct. App. 1969) (refusing to recognize a reasonable mistake of age defense to a defendant charged with selling marijuana to a minor).

these doctrines are consequential exceptions to the usual requirement of desert. They cannot guide action properly and they are unfair.

The moral and legal critiques of pure strict liability are powerful.<sup>123</sup> An agent may be guided by strict liability rules, but not in any way that fairly depends on desert. An agent who fears pure strict liability may simply refrain entirely from engaging in the activity that may create it. Thus, for example, if the potential or current CEO of a supermarket chain fears that he or she will be strictly liable for unsanitary warehouse conditions despite all reasonable efforts to prevent them, the CEO can be guided either never to enter or to exit the industry.<sup>124</sup> The guidance is purely consequential, however, if the underlying conduct is socially acceptable and the defendant has taken all reasonable steps. Strict liability creates hyper-deterrence. The agent can be guided, but only by being placed in an unfair choice situation: engage in socially desirable conduct in a faultless manner and run the risk of criminal blame and punishment anyhow because no one can prevent all accidental harms at reasonable cost, or, refrain entirely from engaging in socially desirable behavior at reasonable cost. This is not the type of action-guidance fair criminal law should provide. Pure strict liability is consequentialism with a vengeance. Furthermore, in most cases, there seems little reason to believe that the consequential benefits of easing the prosecution's usual burden of proving fault are sufficiently compelling to justify abandoning proof of fault.

The only exception to the claim that strict liability ignores genuine culpability might be the few cases in which hyper-deterrence can produce particularly valuable protection with little cost. For example, consider statutory rape laws that make age a strict liability element because the legislature wishes to prevent adults from having even consensual sexual contact with vulnerable minors who have not reached the age of consent. Unlike the CEO in the previous example, the potential statutory rapist can virtually always prevent a violation with little effort by obtaining assurance that the partner is of age. Viewed this way, however, statutory rape sounds like a crime of negligence or even recklessness. Every adult knows or should know that relatively costless steps must be taken to assure oneself about the potential partner's age. On the other hand, if the defendant did take such steps but was nonetheless mistaken—say, asking to see a birth certificate that turned out to be an excellent fake—the defendant would still be liable. In sum, pure strict liability for statutory rape approaches a guidance justification that seems linked to fault, but it is still objectionable. The better solution is to make the standard of care

---

123. See, e.g., MODEL PENAL CODE § 2.05 explanatory note (1962); Reference re Section 94(2) of Motor Vehicle Act, [1985] 2 S.C.R. 486 (holding that it is unconstitutional under the Canadian Charter of Rights and Freedoms to imprison a defendant for a crime of absolute liability); Henry M. Hart, Jr., *The Aims of the Criminal Law*, 23 L. & CONTEMP. PROBS. 401, 422–25 (1958).

124. See *United States v. Park*, 421 U.S. 658, 660–67, 670–73 (1975).



for seeking the requisite assurance very high, but to allow a defense if the defendant took all reasonable steps to avoid conviction. Guidance would then be tied to desert.

Bump-up strict liability may appear to have an action-guiding, desert justification. In these cases, the underlying conduct is criminally culpable and everyone knows that his or her actions may have consequences beyond those initially contemplated. The guidance function directs the agent not to commit a crime in the first place and it therefore may seem fair to hold agents liable for all the untoward results of crimes that they should not have committed. This justification does not survive scrutiny according to the desert-responsive, action-guiding account, however. Bump-up liability, too, is resolutely consequential.

Each crime carries a penalty that in principle is proportionate to desert, which in turn depends in large part on the harm risked by the prohibited conduct.<sup>125</sup> Consequently, the standard rule concerning mens rea is that conviction of every crime should rest on proof beyond a reasonable doubt of its own elements, including the mens rea. The mens rea of one offense cannot be used as a substitute for the mens rea of a further offense that includes a harm that the first offense produced. The legislature has made a judgment that the mens rea of the second crime is necessary and sufficient to achieve proportionate blame and punishment for that crime.

For example, in a famous English case, *Regina v. Cunningham*, the defendant ripped a gas meter from the wall, intending to steal the money in the meter.<sup>126</sup> The defendant was plainly guilty of larceny. Unfortunately, ripping the gas meter from the wall caused a gas leak that exposed the victim to the gas. The defendant was charged with a statutory offense against the person crime that included as an element causing a noxious thing to be taken so as to endanger another person. The court held that the defendant could not be convicted of the second offense just because he was doing something wrong—stealing—that caused the second harm. That would be strict liability for the second offense. The legislature had decided that stealing is one harm that requires fault and endangering another person is another harm that also requires fault. The mens rea for each had to be proven independently to avoid an unfair conviction for the second.

From the standpoint of action-guidance, bump-up strict liability makes little sense. The agent who intentionally commits one offense can be guided by the rule defining it. Proper culpability is thus for the first offense. The further consequences, however, are just bad luck, unless the defendant contemplated the second offense and thus could have been

---

125. The penalty should also be sufficient to provide adequate deterrence. It need not provide perfect deterrence, however. We do not want to reduce the risk of crime to zero, even for serious crime, because doing so would be too costly.

126. *Regina v. Cunningham*, [1957] 2 Q.B. 396, 398 (C.A. 1957).

guided by the rule that defined it. In the latter case, the prosecution should be required to prove all the elements of the second offense, including the mens rea, rather than relying on bump-up strict liability. In *Cunningham*, the court held that the second crime required the mens rea of recklessness. It is quite possible that when the defendant ripped the gas meter from the wall he consciously disregarded a substantial and unjustifiable risk that he would expose a neighbor to danger. Conviction for any lesser mens rea deprives the guidance function of its culpability rationale.

Suppose that some offenses routinely lead to certain unlucky further consequences. Perhaps bump-up strict liability can properly guide the agent in these cases. If the further consequences are sufficiently likely, however, the legislature should enhance the penalty for those offenses to signal that they are more dangerous than they seem, and that those who commit them are more culpable. This would guide the primary conduct more precisely and would be fairer to the defendant than bump-up strict liability. Further, when the further consequences do occur, the prosecution should usually have little trouble proving recklessness or knowledge concerning those consequences. And if the prosecution cannot prove fault for the further consequences because the defendant took special care in committing the offense, bump-up strict liability would be decidedly unfair. In this case, the defendant was guided properly by the enhanced penalty for the primary crime and should have the benefit of his or her partial virtue.

If, as is more commonly the case, the further consequences are not highly statistically likely, bump-up strict liability a fortiori cannot guide action properly and is unfair. Most defendants will commit the primary crime without awareness that the further consequences are likely because, as a statistical matter, they *are* unlikely. One could argue that the defendant should be aware of unlikely consequences. But negligence is a generally disfavored mens rea in criminal law, and why should a defendant be culpable for failing to be aware of small risks? In principle, all crimes can create the risk of unlucky, further harmful consequences. Without proof of mens rea for such consequences, however, the potential agents will not be guided. They will simply be part of a lottery that is not dependent on their culpability.

Let us consider some concrete examples, beginning with felony murder based on the underlying felony of armed robbery. I choose armed robbery because it is one of the felonies that is most inherently dangerous to life. Thus, it presents the most sympathetic case for bump-up strict liability. Consider the increasing progression of penalties for larceny from the person, robbery, and armed robbery. The latter carries the highest penalty because it ordinarily creates a nontrivial but nonetheless low risk of death or grievous bodily harm in addition to the harm of

larceny.<sup>127</sup> If the legislature thought in general that armed robbery was as culpable and as dangerous as murder, then it would punish them equally. But this is seldom the case. The armed robber who commits the crime within the ordinary range of behavior for that crime is extremely blameworthy, but not as culpable as a murderer. If a death ensues accidentally, then this is already reflected in the high penalty for armed robbery compared to the lesser felonies of simple robbery or grand larceny from the person. If the armed robber is negligent concerning death—and perhaps all are—then potential conviction for *involuntary manslaughter* can guide the robber's conduct and bump-up strict liability for *murder* should not attach.

In the ordinary case of armed robbery, however, the culpability assessment is that this crime does *not* create a sufficient risk of death to warrant equation with the culpability for reckless “depraved heart” murder. Therefore, potential conviction for murder can fairly guide the armed robber's conduct only if the crime is committed in an extraordinarily dangerous manner and the robber was aware that he was creating an exceptionally grave risk of death. Thus, the prosecution should have to prove recklessness and the extraordinarily grave risk. If the armed robber plans to kill during the robbery, or forms the intention while committing the felony, then the prohibition against intentional killing can guide conduct and the prosecution should have to prove intent.

For a final example, consider the causation rule that demands that the defendant must take the victim as the defendant finds him or her.<sup>128</sup> If the defendant intentionally commits a trivial harm, such as a minor battery, but grievous harm ensues because the victim is unusually vulnerable, then the defendant will be held liable for the greater harm. The justification is that the underlying conduct is criminal and everyone knows, of course, that some small percentage of the population has physical vulnerabilities that may create the danger of harms greater than those intended by the wrongdoer.

Although an “assumption of the risk” argument may have validity in tort, it does not have justification in criminal law. The legislature, too, knows of the small risk of special vulnerabilities, but nonetheless imposes generally low penalties for less serious offenses against the person. The risk of more seriously harming the small number of potentially vulnerable victims is simply too low to justify enhanced blame and penalty. When death or grievous bodily harm does unexpectedly occur, this is just bad luck unless the assailant is aware that the victim is vulnerable. Potential victim vulnerability cannot fairly guide action in this case. Once again, punishing for the greater harm is unfairly akin to a punishment-enhancement lottery. The underlying criminal conduct does not risk

---

127. See KADISH & SCHULHOFER, *supra* note 88, at 454 n.5.

128. HART & HONORÉ, *supra* note 101, at 79–80. The discussion of this rule is also relevant to the previous discussion of causation. See *supra* Part IV.C.

great harm. The defendant should be punished for no more, even if greater harm unfortunately occurs.

Bump-up strict liability thus abandons the usual rule that the prosecution must individually prove all the elements of every crime charged. The traditional rule is sensible and fair because it focuses *ex ante* on the defendant's action and thus provides maximum action-guidance rationalized by desert. The speculative marginal deterrence provided by bump-up strict liability is achieved only at the cost of over-punishing underlying criminal behavior that unluckily produces further results.

#### F. *Justification*

Justifications offer a rich source of theoretical and doctrinal dispute that the action-guiding approach might help resolve. This section tries only to suggest the virtues of the approach and does not try to consider all aspects of the theory of justification. I claim that the criminal law would be more rational and just if it adopted a subjectivist approach to justifications that focused on the agent's *ex ante* beliefs rather than on *ex post* understandings of the actual circumstances that obtained.

Consider the following examples, which have divided criminal law scholars. What is the appropriate liability of an agent who kills the victim for no justifiable reason, but who, unbeknownst to the agent, in fact was about to be killed by the victim, again for no justifiable reason?<sup>129</sup> Is the agent justified in killing the victim? If so, what, if any, is the agent's criminal liability? Objectivists argue that the agent was objectively justified, despite what she may have subjectively believed and intended. Therefore, the socially desirable result obtained and the agent's deadly force was legally justifiable. Objectivists argue further that the situation is one of impossible attempt and, therefore, the agent should be guilty of attempted murder. In contrast, subjectivists focus on the agent's beliefs and intentions and would convict the defendant of murder.

Should an agent be justified or excused for using self-defensive force when the agent honestly and reasonably *but mistakenly* believes in the need to use such force? Some argue that the agent has done wrong, but since the agent's belief was reasonable but wrong, the agent should be excused. Others think the agent did the right thing, or at least acted permissibly, and was therefore justified. Can two agents using self-defensive force against each other both be justified?<sup>130</sup> Some believe that two agents intending to harm the other cannot both be doing the right thing. Others find such a conclusion unobjectionable. Such cases—and they can be endlessly proliferated—tend to elicit strong but conflicting intuitive responses.

---

129. See, e.g., PAUL H. ROBINSON, *STRUCTURE AND FUNCTION IN CRIMINAL LAW* 108–11 (1997).

130. See, e.g., Douglas N. Husak, *Conflicts of Justifications*, 18 *LAW & PHIL.* 41, 41–42 (1999).

Conduct that is otherwise criminal is justifiable if in the specific circumstances it is the right or permissible thing to do from the viewpoint of society and the law. Conduct that is otherwise criminal but justifiable under the circumstances always risks causing harm, and this is regrettable. Nevertheless, we still prefer or permit the agent to risk causing the harm because doing so is the socially desirable act under the circumstances. The law does not demand under threat of criminal punishment that a justified agent must act on the legal permission. Failure to act justifiably is not a criminal omission, but it is implicit in the notion of justification that agents in similar circumstances are encouraged to act justifiably. After all, from society's viewpoint, justified conduct is preferred. Other than moral saints, few rational agents will have good reason to reject the legal permission.

Conduct will be justified only if it is objectively right under the circumstances. We require agents to form reasonable beliefs about the factual circumstances, which, if true, would make the conduct objectively right. Agents do not have legal permission to do what they subjectively think is right. We all have a duty to avoid unjustifiable harm, and, consequently, we all have a duty to form reasonable beliefs before acting in a manner that would otherwise be criminal.

The difficulty facing potentially justified agents is that no one is a perfect evaluator of the truth of his or her own beliefs. What beliefs are reasonable, what care morality and the law require that we use in forming our beliefs, depends of course on the interests at stake. When deadly self-defensive force is at stake, for example, we require the agent to be very careful indeed before concluding that the use of such force is permissible. But anyone can be mistaken, even if the agent has exercised all the care that morality and the law could fairly demand in the circumstances. Furthermore, incorrect belief formation is particularly problematic in situations of potential justification, which are almost always anxiety-provoking and require immediate action.

The law can only guide agents as practically reasoning people; it cannot guide states of affairs. Just as King Canute could not command the tides, the law cannot command outcomes.<sup>131</sup> The law's command to do the right thing when harmful conduct is potentially justified is thus addressed *ex ante* to an agent forming beliefs about what the agent has good reason to do. It does not direct the agent to act only if the agent is infallibly certain of the justifying facts. This would be impossible and unfair. The law instead directs the agent to use all due care in forming beliefs and gives the agent permission and even encouragement to cause apparently justified harm if she does so. Law and morality can expect no

---

131. THE CHRONICLE OF HENRY OF HUNTINGDON 199 (Thomas Forester ed., 1853) (describing Canute's [King of England from 1017–1035] alleged attempt to command the tides not to rise and to wet his royal self). It is often thought that Canute believed he was sufficiently powerful to hold back the tides, but in fact he was simply demonstrating the limits of kingly power.

more of imperfect agents who are trying to behave rightly. An agent is therefore justified if she forms reasonable beliefs and acts based on those beliefs.

If the agent's reasonable belief that justification obtains is discovered *ex post* to be mistaken, then that is a cause for great regret, but not for condemnation, guilt, or the conclusion that the person was not a morally responsible agent. The reasonably mistaken person was entirely a morally responsible agent and behaved flawlessly; there has been no wrongdoing. The agent caused unnecessary and regrettable harm, but this does not indicate that the agent violated any moral or legal expectation. Indeed, under similar circumstances, the law would *ex ante* encourage her to do it again. A reasonably formed belief does and should give the agent good reason to act, because reasonably formed beliefs are likely to be true according to our best standards of evaluation and truth.

Cases of reasonably mistaken justification are still cases of justification, not cases in which the agent has done wrong but should be excused. Agents are excused when they are incapable of being guided by the law. By contrast, the reasonably mistaken agent who carefully evaluates a situation and acts only after forming a reasonable belief is entirely rational and has been rationally guided by the requirements of morality and the law.

Paul Robinson argues that treating reasonably mistaken agents as justified rather than excused fails to "articulate the criminal law's commands to those bound by them."<sup>132</sup> He wonders why the law would tell people to engage in objectively unjustified conduct. He claims that the agent's belief is not part of the rule of conduct and should become relevant only in "adjudicating failures to follow the rule [to satisfy the ideal]."<sup>133</sup> If he is right, then the subjectivist approach will undermine the criminal law's proper action-guiding function. But the purpose of law's conduct rule in cases of potential justification is not to satisfy the objective ideal. Rather, it is to encourage agents to act so as to maximize the possibility that the objectively ideal state of affairs will obtain.

Morally ideal conduct is acting as well as one can be expected to act under the circumstances. If the circumstances on a second occasion were *precisely* the same—anyway an impossibility—then the agent would have reason to know the second time that circumstances were deceptive and justification did not obtain. At a sensible level of generality in the description of "similar circumstances," however, agents would have every reason to believe that the first situation was highly unusual and should not guide behavior. Therefore, agents would have every good reason to do it again, would be justified in so doing, and would almost certainly be factually right the second time.

---

132. ROBINSON, *supra* note 129, at 118–21.

133. *Id.* at 118–19.

Focusing on the law's guiding function also explains why the subjectively culpable agent who would be justified if she knew the facts should be guilty of the completed crime and not attempt. In a legal regime in which criminal liability depends solely on the agent's intentional action, there would be no distinction between completed crimes and last-act attempts. But even assuming the distinction, the legal conclusion should be the same. The law cannot guide correct action if the agent is unable or unwilling to get the facts right or is intent on doing wrong.

Making liability depend on objective circumstances that were not part of the agent's practical reasoning also leads to doctrinal oddness. When an agent has no idea that she is justified and thinks that she is doing wrong, we would certainly not encourage her to act the same way in similar circumstances, even if it transpires *ex post* that the agent was objectively justified. To say that the conduct was justified, but then to convict the agent of attempt is similarly strange. And what if the agent failed to complete the intended crime? Does this now qualify as an impossible attempt, and if so, is it a case of *factual* or *legal* impossibility? If the circumstances were as the agent believed them to be, then it certainly would have been the crime of homicide. Thus, the impossibility is factual and the agent is not justified at all. But if one views the situation "objectively," trying to kill the potential victim was the right thing to do, which makes the case one of legal impossibility. This is all most strange. None of these arguments from strangeness uncontroversially settles the question of how we should treat these cases, but an approach that produces such convoluted reasoning is surely suspect and should be replaced, if possible, by a more parsimonious account.

As a final example, there can be no genuine conflict of justification if the law focuses, as it should, on each agent's subjective beliefs, desires, and intentions. Multiple agents who each reasonably believe that his or her conduct is justified should be encouraged to act as they reasonably believe that they should. It is, of course, possible that some or even all the agents in such a situation might not be justified from the God's-eye viewpoint (*sub specie aeternitatis*), but people cannot be expected to be gods. As long as agents form their beliefs with due care and act for the proper reason, each is acting permissibly and *ex ante* should be encouraged so to act. There is no conflict of justifications.

One can complicate the picture by introducing further parties with perfect information and asking what they should do. If all immediate parties are in fact justified by the God's-eye viewpoint, which will seldom be the case, then a tragedy is inevitable and the potential intervener should probably do nothing. Almost certainly, the intervener cannot improve the outcome. Suppose, however, that one of the agents is mistaken and the other is not. For example, assume that a fourth agent observes a third party interfering with a lawful arrest because the third party reasonably but erroneously believes the person being properly ar-

rested is in fact the victim of an unlawful assault. Again, assume that the fourth agent has perfect information, that is, she knows that the third party is acting for the right reasons after reasonably forming the belief that the action is necessary. Both the police officer and the third party are justified, but the latter is mistaken. With perfect information, the fourth agent should, of course, try to produce the result that ex post we would all prefer—namely, that lawful arrests be made successfully and without interference. Thus, the fourth agent, if she intervenes at all, should intervene to help the police officer, but not because the third party is acting wrongly. Rather, she should try to produce the socially optimal state of affairs. If the agent does not have perfect information—as virtually all agents in such situations will not—then she will be justified if she forms a reasonable belief about which party needs assistance. It might be very difficult to form a reasonable belief in such cases because a reasonable agent will know that she lacks crucial information and perhaps should not intervene at all.

I conclude that consistent focus on the reasonableness of a potentially justified agent's subjective beliefs and reasons for action yields a much more logical and normatively desirable system of justification than the alternative, objective approach.

#### *G. Results and Epistemic Value Added*

The admission of results for evidentiary reasons, as evidence of desert, is distinguishable from treating results as having an independent, moral contribution to desert. I have argued that results do not contribute independently to desert. It follows that a result, such as a homicide or the destruction of property, and causation of that result need not be elements of a crime and need not be proven beyond a reasonable doubt. Results should be admissible in evidence, however, if their probative value concerning mens rea or the level of risk created, which are criteria for desert, outweighs the costs of admitting such evidence.<sup>134</sup> Resolving the question of the evidentiary value of results evidence is tangential to the central purposes of this article, but a few observations will be useful.

The criminal law can make two kinds of mistakes concerning desert: it may find nonculpable defendants culpable or culpable defendants more culpable than they really are (convicting the innocent or over-convicting); and it may find culpable defendants nonculpable or less culpable than they really are (acquitting the guilty or under-convicting). Our criminal justice system much prefers the latter to the former error because it recognizes the afflictive harm to a citizen of the stigma and punishment attending conviction. The risk of error or nonpersuasion should therefore be assumed by the prosecution. Consequently, due

---

134. Perhaps it is this practical consideration, rather than desert, that explains why results are considered so important.



process requires that the state must prove all the elements of the crime charged beyond a reasonable doubt.<sup>135</sup>

The question is which type of error results evidence is more likely to promote. I conclude that results may have probative value concerning desert, but far less than it may seem. Moreover, admission of evidence of results in many cases poses substantial danger of erroneous judgments, especially errors unfavorable to the defense when results occur. At the least, the minimal value of results evidence should strengthen the claim that results do not contribute independently to moral desert.

Consider first that the criminal law already acknowledges that *mens rea* can be proven without results evidence. For example, traditional attempt liability requires proof of true purpose to complete the crime, such as the true purpose to kill.<sup>136</sup> Thus, in preparatory attempt cases, the state must prove purpose, the most demanding and culpable *mens rea*, *without* results evidence to help confirm the agent's true purpose. In cases of pure risk creation crimes, such as reckless endangerment, the prosecution also must prove both *mens rea* and sufficient risk creation in the absence of results evidence.

Despite the lack of results evidence to guide fair fault ascription in such cases, few people recommend abandoning preparatory attempt liability or liability for reckless endangerment because we cannot be sufficiently confident about *mens rea* or risk to warrant blame and punishment. Indeed, in cases of preparatory attempts to commit serious crimes, we are willing to impose harsh punishment without results evidence to guide us. Finally, our criminal law is willing to convict a defendant of homicide even if the state cannot produce the dead victim. In other words, the law will accept a verdict beyond a reasonable doubt of intentional homicide in the absence of tangible, physical evidence that the alleged victim is not alive. In general, judges and juries are fully capable of assessing *mens rea* and risk creation without results evidence to guide them.

Even if results do not have independent moral status as contributors to desert, it seems a matter of common sense to conclude that results may sometimes be epistemologically valuable as indicators of both the harm risked and the agent's culpable mental state. As we have seen, without results, the trier of fact can evaluate these issues using evidence of the defendant's action and the surrounding circumstances. The evidentiary issues, then, are the value added by results in evaluating the *mens rea* in last-act attempts and completed crimes and for evaluating both *mens rea* and the degree of risk in fully created risk and realized risk cases.

---

135. *In re Winship*, 397 U.S. 358, 364 (1970).

136. There are rare exceptions. *See, e.g., People v. Thomas*, 729 P.2d 972, 977 (Colo. 1986) (holding that attempted reckless manslaughter is a crime in Colorado).

Assume the existence of cases in which results are considered to be the “gold standard” because the fact finder will have the fullest possible information from which to infer both mental state and degree of risk creation. If results were not admissible, would the criminal justice system be more likely to convict the innocent or to acquit the guilty? Results have the potential to create hindsight bias and thus to be misleading indicators of both risk and mens rea, leading to errors in both directions.<sup>137</sup> Results also have strong emotional influences. People are understandably resentful and viscerally aroused by actual harms, and strong emotions affect the rationality of decision making.<sup>138</sup> Consequently, when results occur, we will tend ex post to overestimate the necessary level of risk and the presence of purpose to cause the result; when they do not, we will tend to underestimate. Thus, if results were not considered, the criminal justice system would be more likely to under-punish by acquitting the guilty or at least finding defendants guilty of a lesser offense. Thus, if excluding results evidence produces errors, it will cause the preferred error in our criminal justice system. If results were considered, then we run the risk of over-punishing (or “over-convicting”) when results do occur, and again, under-punishing when they do not. If admitting results evidence produces errors, then it will cause the disfavored error in cases in which results actually occur.

Although there is cause for concern about the effect of results evidence, the influence of general hindsight bias and emotion in criminal justice decision making is unknown. Moreover, results evidence often appears sufficiently probative to outweigh the potential prejudice. For example, the manner in which death occurred may be highly probative, and it will be impossible to introduce this evidence without the fact finder learning that the victim died. Further, if judges and juries do not learn of the results, they may tend to “complete” the story using their own imaginations, which would create unpredictable effects.<sup>139</sup> In sum, it would be difficult to exclude results evidence, even if admitting it creates dangers. As a practical matter, the potential for bias must be addressed through admission of evidence on this point, much as claims about the

---

137. These types of errors have been amply documented in the psychological literature concerning heuristics and biases. See, e.g., Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, in *JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES* 3, 3–4 (Daniel Kahneman et al. eds., 1982). See generally Jeffrey J. Rachlinski, *A Positive Psychological Theory of Judging in Hindsight*, 65 U. CHI. L. REV. 571, 576–88 (1998) (canvassing the hindsight bias literature and offering a new account).

Much of this research concerns negligence rather than subjective mental states. Other experimental work relevant to hindsight bias applies to a potentially wider range of retrospective blame determinations. See Mark D. Alicke, *Culpable Control and the Psychology of Blame*, 126 PSYCHOL. BULL. 556 (2000) (reporting relevant experimental results, but adopting a theory of blame quite different from the theory presented in the present article).

138. See Cass R. Sunstein, *The Laws of Fear*, 115 HARV. L. REV. 1119, 1139 (2002) (book review).

139. See Linda L. Carli, *Cognitive Reconstruction, Hindsight, and Reactions to Victims and Perpetrators*, 25 PERSONALITY & SOC. PSYCHOL. BULL. 966 (1999).

accuracy of specific eyewitness identifications can be challenged by evidence about the general accuracy of eyewitness testimony.

Finally, if evidence of the occurrence of results is admitted because it is considered probative, should causation doctrine then be used as an evidentiary test to decide when otherwise probative results evidence should be excluded? If the causal chain from the defendant's action to the result is too "unforeseeable" or "remote" or "accidental" to permit liability as an independent, substantive matter, then how can evidence of results in such cases be probative about desert? But one never reaches the causation question, whether treated as an independent substantive matter or as an evidentiary test, unless the defendant's action and mental state satisfy the criteria for the crime charged. What this means, of course, is that in the ordinary course of events, the results would have occurred in an expectable manner. The causation question arises only because there has been deviant intervention *after* the defendant has done everything possible to cause the result. Indeed, even when the causal chain is broken, the defendant's action is still a but-for cause of the result. Thus, if the defendant is a but-for cause and the result would have occurred in the ordinary course of events, why shouldn't the result be admitted if it is probative of culpability, even if a deviant cause intervened to "complete" the result? Result evidence, if otherwise probative, should not be excluded by deviant causal chains because a deviant causal chain does not make it less relevant to the defendant's culpability. This conclusion bolsters the contention that results do not contribute independently to desert.

In conclusion, results evidence probably has only limited additional value for proving mens rea and the amount of risk creation beyond the value of the fullest evidence of the defendant's behavior and the surrounding circumstances. Moreover, results evidence creates the risk of error, especially the disfavored error of over-convicting when the result actually occurs. Nonetheless, such evidence will be admitted and the remedy for the potential dangers must be the rigors of the adversary process.

#### V. THE COUNTERARGUMENT: RESULTS MATTER

It is difficult to justify differential treatment of defendants depending on whether or not results occur because it seems apparent that legal rules and moral considerations can guide only actions. An agent's moral desert and, for the most part, dangerousness are not dependent on results. Nonetheless, the law and ordinary people continue to treat results as if they matter, and some have tried to defend doing so on desert grounds. Professors Michael Moore and Leo Katz, who are both leading criminal law scholars, provide richly detailed arguments for why results

independently contribute to an agent's desert.<sup>140</sup> In the first section of this part, I consider Professor Moore's particularly complete arguments and conclude that they are not persuasive.<sup>141</sup> Then I turn to Professor Katz's characteristically ingenious argument and conclude on both substantive and methodological grounds that it does not make the case.<sup>142</sup> This part next briefly responds to Professor Stephen Perry's brief but intriguing suggestion that results do increase blameworthiness.<sup>143</sup> The next section challenges other answers that have been given. Finally, I turn to practical arguments based on psychology and sociology and conclude that they provide the best consequential reason for why results matter, but that these arguments are still weak.

A. *Michael Moore on the Independent Relation of Results to Moral Desert*

Michael Moore firmly believes that culpability is necessary and sufficient for desert and that culpable attempts are therefore punishable. In contrast, the amount of harm an agent actually causes—what he terms “wrongdoing”—is neither necessary nor sufficient for desert, but nevertheless independently contributes to a culpable offender's moral desert. When they occur, results add to desert. Wrongdoing is therefore “something of a poor relation” to culpability, as Moore puts it.<sup>144</sup> As should be clear from parts II and III, I find the usage of the term “wrongdoing” to mean the amount of harm caused an unfortunate and question-begging locution. An agent does wrong by wrongful action, that is, action that breaches moral or legal expectations. Such wrongful action may or may not cause harm, but is nonetheless wrongful. I would prefer the term “harmdoing” to refer to the actual harm caused, but will continue to use Moore's term in this section. Finally, Moore provides both foundational and nonfoundational arguments for his conclusion. Because he is not a foundationalist and believes that foundationalist and nonfoundationalist arguments merge, I shall discuss only the latter.

How does wrongdoing, the “poor relation,” add to desert? Moore's most powerful argument addresses the standard contention that result luck is distinguishable from luck concerning the antecedent causal variables that influence behavior, and, therefore, we are not responsible for results but we are responsible for our behavior.<sup>145</sup> He argues that “result

---

140. Neither Moore nor Katz endorses the strict liability rules that I criticize, however. See *supra* Part IV.E (discussing strict liability).

141. MOORE, *supra* note 50, at 218–46. Moore's defense of the independent significance of results is so long and detailed that my response also must necessarily be lengthy and detailed to do him justice.

142. Leo Katz, *Why the Successful Assassin Is More Wicked Than the Unsuccessful One*, 88 CAL. L. REV. 791 (2000).

143. Perry, *supra* note 94, at 94–95.

144. *Id.* at 191–93.

145. I have already discussed this contention. See *supra* Part III.

luck” is indistinguishable from “constitutive luck” (luck involving the genetic and experiential fortuities that cause one to have the character that produces potentially harmful conduct), “planning luck” (luck involving fortuities that may intervene to prevent one from forming plans to cause desired potentially harmful conduct), and “execution luck” (luck involving fortuities that prevent the execution of firmly formed intentions to perform potentially harmful conduct).<sup>146</sup> Moore convincingly demonstrates that in a causal world, variables over which the agent had no control exist at every stage, from character to execution. Whether or not one performs a wrongful action is as suffused with luck, and is as beyond our control, as whether results actually occur. If we had a different character, which is largely not up to us, then we wouldn’t be the type of person who would have the desire or form the intention to do wrong. If we have the desire to do wrong, then events over which we have no control may distract us from forming the intention to satisfy that desire. Once we form the intention, events over which we have no control, ranging from physical to psychopathological, may prevent us from executing that intention.

As I suggested, the conclusion that luck, understood in this way, “goes all the way down” is correct as a matter of theoretical reason.<sup>147</sup> But as Moore recognizes, it leads to an unacceptable reductio that those who believe results do not matter will be forced to accept. If causal influences over which we have no control undermine responsibility, then no one is responsible for anything, and this conclusion is not limited to results. Genuine desert does not exist. Because this conclusion is morally unacceptable and does not account for the moral and legal world we inhabit, a morally principled line based on luck cannot be drawn. Moore suggests that because we accept the existence of desert for action despite constitutive, planning, and execution luck, we should be willing to accept the existence of desert for results despite result luck. Moore’s suggestion does not follow, however. Luck may not provide a principled basis to draw the line between moral responsibility for action and moral responsibility for results, but perhaps another principle that does not lead to an unacceptable reductio will.

Moore’s critique of the luck argument depends on that argument’s confusion of the distinction between a mechanistic-causal account of behavior and a practical reasoning account.<sup>148</sup> As a good compatibilist, Moore knows that the explanatory causal stories of theoretical reason are not the basis for responsibility and desert. They could not be for just

---

146. MOORE, *supra* note 50, at 233–46.

147. See *supra* Part III. On the distinction between theoretical and practical reason, see BOK, *supra* note 5, at 52–91 (employing the distinction to ground a compatibilist account of moral responsibility).

148. See MOORE, *supra* note 50, at 610–36 (explaining the law’s view of the person as a practical reasoner); Morse, *supra* note 28, at 337–40 (distinguishing the two conceptualizations or views of the person).

the reasons he gives. Causation is a metaphysically seamless web. In a compatibilist account of the type he endorses, responsibility and desert are moral concepts implicit in practical reason. Compatibilists (and libertarians) believe that human beings are capable of rational deliberation, that our deliberations affect our actions, and that our actions potentially affect the world. For compatibilists, agents do not need contra-causal freedom to deserve praise and blame, or punishment and reward, for their actions. It is sufficient if they act intentionally, without compulsion, and with a general capacity for rationality.<sup>149</sup> The only form of control a responsible agent needs is the general capacity to be guided by reason, a capacity most adults possess in ample measure.<sup>150</sup> Indeed, successful human interaction would be impossible without it. Thus, compatibilists have good reason to “draw the line” at human action because only action can be guided by reason and not because action is free of the causal forces of the universe—of “luck.” There is always a causal story, but results cannot be guided by reason. The potential for the law to guide people by reason is a good justification to hold people morally responsible for action but not for results. Moore needs other arguments, unrelated to luck, to demonstrate that results should matter to desert.

The following discussion has two parts. The first considers Moore’s moral-coherentist methodology as it applies to the question of results. I conclude that the success of his coherentist strategy for proving that results matter depends entirely on already accepting that results do matter. Thus, Moore initially begs the question. I then turn to his attempt to avoid begging the question. Moore tries to offer a series of independent arguments about why results should matter to desert that are based on broad classes of judgments and experiences and that employ emotional reactions as guides to moral truth. I conclude that none of these independent arguments ultimately persuades that results should matter to desert.

### 1. *Moral Coherentism and Results*

Moore’s general approach to moral justification is a “coherence,” nonfoundational view. First principles are not self-evident, but must be inferred from more particular judgments that for these purposes must be accepted as true.<sup>151</sup> To argue for such nonfoundational first principles

---

149. Moore essentially agrees with this. MOORE, *supra* note 50, at 217 (“By our ordinary notions, we control our choices whenever such choices are not subject to threats or other coercion and when we have enough information to make them.”). I would use different terminology, but the idea is the same.

150. Whether an agent has this capacity is a function of variables over which the agent has no control, but this is of no consequence. For example, extreme developmental disability or severe mental disorder may block or undermine the capacity. We excuse agents suffering such disabilities because they lack the capacity for rationality, not because they have been caused to be as they are. Causation is not a criterion for responsibility or for excuse.

151. See MOORE, *supra* note 50, at 225.

will accordingly be to present a broad range of intuitively compelling particular judgments and argue that such judgments are best explained by the general moral truth to be justified.<sup>152</sup>

Moore claims that the best explanation of many of our particular judgments is that “[the] wrongdoing [result] independently determines the extent of our just deserts.”<sup>153</sup> For the purpose of my argument, I will accept the validity of this form of moral justification. Those who do not accept it will inevitably find this part of Moore’s arguments unpersuasive.

Moore’s coherentist defense of the contribution of results to desert begins with the observation that in particular cases when harm occurs, “many” intuitively believe that increased punishment is deserved.<sup>154</sup> He asserts—wrongly, I believe—that this belief is shared even by those who hold the “standard educated view” that results do not matter, but he freely concedes that those who do not share this intuition will not be persuaded by this line of argument. Moore then claims that the general moral principle that best explains the observation is that wrongdoing independently determines the extent of just deserts.

Moore canvasses four objections to this view. The first is that the process of justification is circular: the particular judgments are in fact indistinguishable from the general principle. He rejects this objection because particular judgments are not logically general. They apply only to the specific circumstances of each individual case, whereas the general principle applies across all similar cases. This may be formally true, but as the next response to his argument suggests, it may also be trivially true.

The second objection is that the general principle that results contribute to desert is prior to and generates the particular judgments, rather than the reverse. Moore claims, however, that even if this is true, it does not affect the justificatory power of the singular judgments. The singular judgment, he says, is simply, “A deserves more punishment,” and not, “A deserves more because results matter.” According to Moore, why A deserves more is an open question that is answered by an inductive process that attempts to determine which characteristics of agents who simply deserve more are the characteristics that in fact increase desert.<sup>155</sup> When we engage in this inductive process, we “discover” that results do matter to desert, but, for example, that eye color does not.

The difficulty with Moore’s answer is that judgments about desert are moral judgments and can only be made according to some moral the-

---

152. *Id.*

153. *Id.* at 226.

154. Moore sometimes thinks that only the judgments of morally and psychologically well-formed and well-informed people should count. *See, e.g., id.* at 231. But this surely begs the question. In any case, the “many” count only if they are well-formed and well-informed. In light of this assumption, Moore’s conclusion that results should matter is odd because he calls the view that results do not matter “the standard educated view.” *Id.* at 193–94.

155. *Id.* at 226–27.

ory. To the best of our knowledge about how the brain and sensory apparatus work, there is no moral sensor that reads off moral truths from brute facts. Whether the application of a moral principle is correct has truth criteria that will depend on facts about the world, but whether the principle itself is justified is a matter of practical reason. For example, it is true that some actions cause harm and some do not and it is true that most people believe that this difference matters morally. But these truths give us no independent reason to consider results morally significant unless one believes in a relatively infallible moral sensor that gives us reason to trust our “brute” reactions to events. There is no simple judgment, unaffected by prior theory, whether explicit or implicit, that some people deserve more.

As a matter of psychology, perhaps we may, without the mediation of a theory, simply have certain emotional reactions when harms occur. For example, we may be angrier or more resentful when harms do occur, but having such feelings is not per se a moral judgment. Indeed, we are more inclined to have such emotional reactions because we already have a moral theory that tells us that we *should* be angrier. When we do explicitly reach moral conclusions such as, “A deserves more punishment,” it is always because, at least implicitly, we already accept a principle that some characteristic justifies that conclusion. The general principle is implicit in the singular judgment rather than generated by it. We do not simply and atheoretically judge that “A deserves more.”<sup>156</sup> Moore ultimately recognizes the force of this difficulty and tries to turn to fresh arguments that are independent of and disengaged from widely held theories.<sup>157</sup> Before addressing these arguments, however, I shall briefly respond to the other two objections to his coherence methodology.

The third objection Moore answers is that his coherence methodology is just a heaping up of untutored intuitions that can tell us a great deal about our moral intuitions, but that cannot tell us if those intuitions are justified. The methodology is allegedly too subjective to be an adequate guide to correct moral understanding. With Moore, I assume that morality can be objective.<sup>158</sup> Moore rightly claims that our moral judgments are about objective features of the world and not about subjective mental states. Thus, there are truth criteria for moral judgments such as, “A deserves more punishment,” and our moral judgments will be caused by features of the world. But the objectively verifiable features of the world that move us morally are not abstractions, such as desert, that are

---

156. See MATT MATRAVERS, *JUSTICE AND PUNISHMENT: THE RATIONALE OF COERCION* 81–87 (2000) (reaching—for many of the same methodological reasons I give—a similar conclusion about Moore’s justification of the first principle of desert).

157. See MOORE, *supra* note 50, at 228–29.

158. I do not mean that to be objective one must subscribe to Moore’s theory of metaphysical moral realism. I mean only that one must believe that moral concepts can be justified by reason. See SIMON BLACKBURN, *RULING PASSIONS* 279–310 (1998); Ronald Dworkin, *Objectivity and Truth: You’d Better Believe It*, 25 *PHIL. & PUB. AFF.* 88–89 (1996).



*intrinsic* features of states of affairs or agents. Which features move us is determined by moral principles, such as desert and responsibility, which are practical concepts justified by practical reason. These practical concepts have criteria that are factual properties of agents or events, such as whether the agent had the general capacity for rationality or whether harm occurred. If the criteria are met (or not), then we will tend to conclude that the moral principle is applicable (or not). These singular moral judgments thus depend on a *prior* general moral principle and cannot themselves be the foundations for that principle. So, for example, it is certainly true that most people will respond with judgments of increased desert if, and only if, harm has actually occurred. They can be quite objective about this judgment that a culpable harmdoer deserves more punishment, but only because they already believe that harm contributes to desert.

The final objection to coherence methodology is that it has so often been wrong, even when our judgments have been most uniform and coherent. It is surely true that moral mistakes have been made, but it does not follow that coherence is fatally flawed. Assuming again that morality can be objective, Moore properly answers that principles that seem well-justified may be radically false and are always subject to rejection or revision by better reasons.

Thus, the real question is whether a moral principle is justified by reason. To be more specific, what reason is there—independent of people's reactions to results—to believe that results *should* matter? Moore has not yet answered this question. He says the best way to do this without begging or trivializing the question is to broaden the class of judgments and experiences best explained by the principle that results matter. Let us turn to his attempt to do this.

## 2. *Emotions, Experience, and Desert*

Moore's first, allegedly theory-independent observation is that most people more deeply resent culpable agents who cause harmful results than culpable agents who try but fail to do so. He rightly rejects the unrefined uses of this psychological fact about emotional reactions, but argues that such differential resentment is evidence of the truth of the judgments to which such feelings lead.<sup>159</sup> Moore is here explicitly relying on his view that emotions sometimes respond to objective, intrinsic *moral* features in the world and that the virtue of feeling an appropriate emotion is the best criterion of whether it is epistemically reliable as a guide to moral truth. Assuming that this view in general is correct,<sup>160</sup> the question is whether feeling more resentment when the result occurs is

159. See MOORE, *supra* note 50, at 229.

160. Moore argues for this view elsewhere. *Id.* at 127–38. I, of course, reject this version of meta-physical moral realism, as does Matravers. MATRAVERS, *supra* note 156, at 81–87.

morally virtuous and thus an indication that it is morally true that results matter to desert. Moore argues that resentment can be “virtuous [when] it expresses outrage at the unnecessary suffering of victims of wrongdoing and [when] it expresses hatred of the flouting of morality such infliction of suffering involves.”<sup>161</sup> Such resentment is some evidence for the truth of the consequent judgment that results matter.<sup>162</sup> This conclusion begs the question, however, without independent reason to believe that more resentment is in fact morally justifiable.

Emotions, like cognitive judgments, are propositional.<sup>163</sup> They are about something; they take objects. Emotions are not simply unpropositional feelings—such as the feeling of being cold—or atheoretic reactions to brute facts. If emotions unmediatedly led to judgments, then we would have good reason to distrust those judgments as morally justifiable. The emotional reactions we experience depend upon the meaning of the situation to which we respond. Greater resentment is only morally virtuous if results do matter. Differential resentment is hard to justify, however. The agent who tries but fails to cause harm has fully flouted the moral expectation that he or she will not act so as to risk that harm. Morality, once again, can only guide action and not states of affairs. To the extent that it is ever virtuous to feel resentment against agents who flout moral expectations, we are justified in resenting equally agents who try successfully and those who fail. I agree that when harmful results occur, we are justified in feeling greater negative emotion, but I suggest that the justified emotions are profound regret or sorrow and compassion for the victim, rather than greater resentment toward an agent who is no more culpable than one who failed. Greater resentment may be “natural” in such cases, but it is, I claim, a primitive emotion, and not a virtuous feeling that is a good guide to moral truth. Finally, how do we know that the emotion is a heuristic guide to the moral truth unless we already know what the moral truth is. And if we know it already, we do not need the guide.

Moore’s next contention is about first-person experience. He observes—again accurately—that many people feel greater guilt when they cause harmful results than when they fail to do so, and thus we may infer the truth of the judgment to which such a feeling leads—that they *are* guiltier.<sup>164</sup> People often have “unjustified” emotions, however, among

---

161. MOORE, *supra* note 50, at 230.

162. *Id.*

163. Moore accepts this, as indeed he must. The literature on this point is enormous and controversial, but the propositional/evaluative view of emotion has long been the dominant paradigm. See PAUL E. GRIFFITHS, *WHAT EMOTIONS REALLY ARE* 2 (1997). There are many representative examples of this approach. See, e.g., PATRICIA S. GREENSPAN, *EMOTIONS AND REASONS: AN INQUIRY INTO EMOTIONAL JUSTIFICATION* (1988).

164. See MOORE, *supra* note 50, at 230–31. George Fletcher agrees that those who succeed have more reason to feel guilt than those who fail. GEORGE FLETCHER, *BASIC CONCEPTS OF CRIMINAL LAW* 174 (1998).

which irrational guilt is a prime culprit.<sup>165</sup> More important, once again, is the basic question of whether the emotion leads to the judgment or whether the judgment leads to the emotion. But to avoid begging *that* question, let us leave it open and turn directly to the question of whether greater guilt is virtuous or unwarranted. We ought indisputably to feel guilt for breaching a moral expectation and thereby causing unnecessary suffering, but should we feel any less guilt if we similarly breach but fail to cause harm? Moore claims that not only do we feel more guilt when bad results occur, but that the guilt has a different quality. He asserts that the guilty reaction to “culpable attemptings or riskings . . . has a kind of narcissistic preoccupation to it, a kind of keeping-our-own-moral-house-in-order flavour.”<sup>166</sup> We are sick at what we have done, he says, and a guilt that reflects a concern for others is more virtuous than a guilt concerned only with our own moral health.

To begin, I have no idea how Moore knows that guilt in the absence of results is narcissistic, nor how one could justify this judgment without already accepting that attempts and risks are less deserving or somehow deserving of a *different* type or degree of guilt. Moore accepts that one who risks or attempts is *as* culpable as one who causes harmful results, so why shouldn't the attempter feel equal levels of justified, virtuous guilt? I agree that one should feel sick when one has culpably caused harmful results, but the more justifiable reason is empathy and compassion for another suffering person, rather than greater guilt. Moore is also right that people typically feel lesser guilt and a sense of relief for failure, but I claim that they have no right to let themselves off the moral hook so easily. A vital moral point of guilt feelings is to guide future behavior by indicating when we have behaved badly. Therefore, it is our actions that are at issue, not results. Moreover, as Moore recognizes, the relief following failure is not justified by any aspect of the attempter's conduct.<sup>167</sup> The person's conduct is equally homicidal, whether or not death occurs. Moore says, however, that the person has genuinely escaped the moral guilt of being a killer, a particularly heinous form of wrongdoing. I disagree. He has escaped the pain and regret of causing a death, but he does not deserve to escape any of the guilt for equally homicidal conduct.

---

165. See WALLACE, *supra* note 57, at 40–50.

166. See MOORE, *supra* note 50, at 231.

167. *Id.* Moore loads the deck by using as his example the seriously intoxicated driver who avoids hitting a child. He has chosen perhaps our culture's most common life-endangering conduct. Because so many people have driven while intoxicated without killing, it will be a common psychological reaction that they are not as culpable as actual killers. And, in fact, we are quite unclear about the degree of moral breach the behavior exhibits in general. Driving intoxicated is always a breach of a moral expectation, but how often is it seriously life endangering? If we thought that it always was, I contend that we would have far more serious penalties for the activity *per se*. Thus Moore has chosen an example for which, strangely enough, in some cases the unlucky killer may feel more guilt than is warranted because in many cases the conduct may not be as risky *ex ante* as we think *ex post* when harmful results occur. Finally, one might challenge whether relief always accompanies failure. For example, those who intend harm may feel frustration or anger and may not feel any relief at all.

We feel forever psychologically tied to a victim we have in fact harmed and rarely feel ties to potential victims, but how we should feel morally does not follow from these natural psychological reactions. Consider a case in which we cause a death entirely accidentally or justifiably. I submit that most people would feel profound regret their entire lives, even though they had done nothing wrong.<sup>168</sup> Furthermore, many people would feel great guilt. According to Moore, causing harm contributes to desert independently only if the agent acted culpably, but if emotions are a guide, shouldn't we conclude that causing harm always contributes independently to desert, even in the absence of wrongful conduct?<sup>169</sup> Moore would have to say—and I would agree—that the guilt is unwarranted in such cases; there is no desert and only regret is justified. But we agree about this conclusion only because we agree about the moral theory that guides it. One cannot justifiably be as opportunistic about emotional heuristics as Moore's theory appears to be. He needs an argument about *why* we are justified in feeling guiltier if results occur. This argument he has not yet produced.

Moore's last example of the kind of experience that suggests that results do matter is the "forward-looking experience of choosing what to do."<sup>170</sup> Moore points to the many moral dilemmas we all face and says that we do care what consequences result from our actions. We know that it is not enough simply to decide in the proper manner what to do; results count. Moore is right about this, of course. We all hope desperately for good outcomes. Action matters precisely because action has the potential to affect the world around us for good or for ill. Results matter *ex ante* because their potential occurrence is action guiding. But as Moore recognizes, once one has acted, culpability is complete and "[i]t is only the amount of . . . wrongdoing [harm] that is hostage to the future."<sup>171</sup> He argues that the independent significance of the harm is attested to by our continuing concern for results, even when we have acted as well as we can. Again, he is right that we do and should have continuing concern. We care about our actions primarily because they can cause results and agents morally should hope that good results occur and bad results do not. We are terribly disappointed and feel profound regret if we fail to achieve what we intentionally and properly tried to accomplish and we feel great joy if we succeed. Results clearly have significance *ex ante* and *ex post*, but it does not follow that our desert is independently

---

168. Bernard Williams terms this "agent-regret." BERNARD WILLIAMS, *Moral Luck*, in *MORAL LUCK* 20, 27–31 (1981).

169. It is possible, of course, that morality might not be symmetrical here. The amount of harm might contribute to desert when an agent breaches a duty, but not when an agent acts entirely innocently. It is true as a matter of sociological observation that we tend to suspend reactive attitudes of anger and resentment when we believe that a harmdoer acted innocently, and that our reactive attitudes to breach of duty are in fact proportionate to the amount of harm done. But again, this is a sociological observation, not a moral justification.

170. See MOORE, *supra* note 50, at 232.

171. *Id.* at 233.

attributable in part to results. I suggest yet again that an agent who behaved properly deserves full moral credit; an agent who breached a moral obligation deserves full moral condemnation. Until we are provided an argument about why (undeniably significant) states of affairs that are not capable of being guided by reason should be tied to desert, we are entitled to conclude that results do not contribute independently to desert.

In sum, Moore's argument from emotional epistemology depends on already accepting the conclusion that results do contribute independently to desert. People would not otherwise have the emotional reactions he correctly describes. But that leaves open the question whether the conclusion that results matter is justifiable. Moore powerfully demonstrates that results have psychological and moral significance, but he does not demonstrate that results contribute independently to an agent's desert.

Moore often constructs an argument for his preferred conclusion in the following manner. He identifies the counter-arguments and demonstrates that none succeeds. He then draws the conclusion that his position is probably correct.<sup>172</sup> Although this conclusion is not strictly logically entailed—he might be wrong and the opposite position may be supportable by a convincing argument not yet developed—it is not implausible, especially in areas in which much work already has been done. If my responses to Moore are convincing, then he cannot object if I draw the conclusion that results do not matter to desert.

### B. *Leo Katz and Trying Transitivity*

If *A* deserves more punishment than *B*, and *B* deserves more punishment than *C*, it must be true that *A* deserves more punishment than *C*. Right? Well, that is Professor Leo Katz's argument.<sup>173</sup> It would be correct if each of the individual desert claims he makes was independently correct, but that is where the argument goes astray.<sup>174</sup> For the purpose of my argument, I shall accept the validity of Katz's methodology of reaching moral conclusions using intuitions produced by largely fanciful hypotheticals.<sup>175</sup> My analysis of Katz's hypotheticals will assume with him

---

172. See, e.g., Michael S. Moore, *Moral Reality Revisited*, 90 MICH. L. REV. 2424, 2491–533 (1992) (arguing that all the objections to metaphysical moral reality are incorrect, and concluding, consequently, that metaphysical moral reality is the most convincing metaethical theory).

173. Katz, *supra* note 142, at 800–06.

174. Paul Robinson has written an excellent critique of Katz's position on both substantive and methodological grounds. Paul H. Robinson, *Some Doubts About Argument by Hypothetical*, 88 CAL. L. REV. 813 (2000). I agree with virtually all of Robinson's analysis and will not repeat the points he makes.

175. I should say, however, that I often have no intuitive response to the most incredible hypotheticals. Moreover, I wonder how intuitions generated by fanciful hypotheticals can be valid when the hypothetical is so distant from the ordinary experiences that are the source of our intuitions. My general approach to using fanciful hypotheticals and the intuitions they generate as guides to truth is

that last-act attempts and successes are morally indistinguishable.<sup>176</sup> I make this assumption because I think that it is correct; Katz properly makes it to avoid begging the question. Therefore, in what follows I shall refer to both a last-act homicide attempt and a successful homicide as instances of “Culpable Homicidal Endangerment” (CHE). In addition, CHEs can be graded according to the mental state with which the agent acts to risk the potential victim’s death. For example, agents who homicidally endanger with intent to kill engage in more culpable behavior than agents who homicidally endanger only with conscious awareness that their behavior risks death. My last assumption to ground the analysis of Katz’s argument, which he apparently shares, is that desert is proportionate to the amount of culpable conduct the agent performs.

Katz’s set-up of the hypotheticals is both ingenious and amusing. He creates a series of pair-wise scenarios in which various assassins do multifarious dastardly deeds. For the sake of simplicity, I shall loosely follow Katz’s summary.<sup>177</sup>

Assassin 1 (A1): Poisons five people with the intent to kill them and all five die.

Assassin 2 (A2): Decides to poison five people with the intent to kill them, knowing (miraculously, but hey) that he can save their lives by killing and carving up a sixth victim, whose organs will (miraculously, but hey) somehow and without fail be used to save the five. A2 goes sufficiently far with his poisoning scheme to qualify for a preparatory attempt, but shortly *before* administering the poison, decides that he will save the five after poisoning them. A2 then poisons the five, kills the sixth, and saves the five.

Assassin 3 (A3): Poisons five people with the intent to kill them, knowing with complete certainty that he can save the five by killing the sixth. Shortly *after* poisoning the five and before they die, A3 decides to save the five, kills the sixth, and saves the five.

Assassin 4 (A4): Poisons five people with the intent to kill them, but the poison is unforeseeably weak and none of the five dies.

Katz asks us not to do the usual comparison of A1 with A4, which are both cases of five CHEs, because one’s intuitive response will be based on one’s prior theory about whether results matter. Instead, he first compares A1 to A2, which he characterizes, respectively, as an assassin who hits and an assassin who turns a hit into a miss. Then we must compare A3 to A4. Katz claims that “it seems intuitively evident” that A1 is “worse”<sup>178</sup> than A2 and that A3 is worse than A4. Then, Katz

---

this: Used in moderation, I have no objection; used immoderately, such hypotheticals are unlikely to do anything other than tie one’s brain up in knots.

176. Katz, *supra* note 142, at 802.

177. *Id.* at 806.

178. “Worse” is an unfortunately vague word. It may refer to a judgment of an agent’s overall character, or it may refer to a judgment of total culpability for action. I will assume that the latter meaning is intended.

claims,  $A2=A3$  ( $A2/3$ ). If so,  $A1>A2/3$  and  $A2/3>A4$ , and it follows that  $A1>A4$ . That is, the successful assassin of five is more culpable and deserves more punishment than the unsuccessful assassin of five.

All would agree that  $A1$ , who is guilty of five CHEs, is more culpable than  $A2$ , who is guilty of only one CHE. And all would agree that  $A3$ , who is guilty of six CHEs, is more culpable than  $A4$ , who is guilty of only five. But is it true that  $A2$  and  $A3$  are *equally* culpable, an assumption that drives the entire argument? I think not. Katz thinks that the slight difference between when  $A2$  and  $A3$  form the intent to save, that is, just before versus just after poisoning, cannot make much moral difference. For the action-guiding approach, however, concurrence is crucial. When the assassin decides to save the five makes all the difference.

$A3$  poisons the five intending to kill them. These are last-act attempts and clearly five CHEs are performed at the moment of poisoning. Sure,  $A3$  knows that he can save the five by later killing the sixth, but so what? When he poisons them he intends to kill them. After the poisoning occurs, and staying within the strict confines of Katz's own hypothetical, nobody knows what thoughts will pop into  $A3$ 's head, what feelings will intervene, or what unforeseeable distractions will occur. Maybe he will not form the intent to save the five. At the time he poisons the five with the intent to kill,  $A3$  has no intent to save them and cannot possibly be sure that he will form and act on the intent to save. Thus, in terms of action guiding, there are six CHEs (five last-act attempts and one killing). Further, let us relax just a tiny bit the heroic assumption that at the time of poisoning  $A3$  can know with complete certainty that he can save the five. For example, assume, as Katz does, that if  $A3$  carves up the sixth, the five can surely be saved, but it is not certain that he can carve up the sixth. Suppose someone unforeseeably steals  $A3$ 's scalpel, thus preventing him from committing a last-act attempt to kill the sixth. In that case,  $A3$  is guilty of five CHEs and one preparatory attempt homicide (assuming that preparatory attempts are criminalized).

How does  $A3$ 's renunciation of the intent to kill the five affect his desert? Does it make him less culpable than  $A1$ ?<sup>179</sup> Again, I think not. As we have seen, the renunciation defense is based upon primarily consequential considerations.<sup>180</sup> The law wishes to give agents an incentive not to cause the prohibited harm, and perhaps renunciation indicates that the agent is not so dangerous after all and that no incapacitation is needed. Neither of these considerations bears on desert for the prior conduct, although the incentive effect is fully consistent with the action-guiding function.  $A3$  is fully culpable for six CHEs.

---

179. Katz does not make this argument in the paper, but does so in a personal communication with Paul Robinson. Robinson, *supra* note 174, at 817–19 n.13. Robinson concludes that the major problem with the argument is that we have no firm intuitions about renunciation, and it is unclear how much mitigation is warranted by it. *Id.*

180. See *supra* Part IV.A.4 (discussing the renunciation or abandonment defense to attempt liability).

Now consider *A2*, who does not poison with the intent to kill because, by hypothesis, when he poisons the five he intends to save them and knows with complete certainty that he can save them by killing the sixth. Under the action-guiding view, he is guilty of five intentional poisonings, however that is criminalized, and only one CHE of the sixth. *A2* has committed vastly fewer CHEs and is considerably less culpable than *A3*, who committed six (and *A1* and *A4*, both of whom committed five). Suppose we relax again the “sure save” assumption. That is, suppose either that *A2* might change his mind and abandon the scheme to save or again that someone steals the scalpel. In that case, *A2* would have committed five CHEs, but with a less culpable mental state—either recklessness or negligence concerning the death of the five—than the intent to kill that all the others manifest. Even then, *A2* is still the least culpable because five reckless CHEs are less culpable than five intentional CHEs.

If we focus on action alone as the appropriate basis for desert and leave luck out of it, as Katz does, then *A3* is the most culpable; *A1* and *A4* are equally culpable, but less than *A3*; *A2* is the least culpable. Katz has not yet demonstrated that those who succeed are more culpable and deserve more blame and punishment than those who commit last-act attempts. Katz can, of course, continue endlessly proliferating ever more clever (and fanciful) hypotheticals to drive intuitions toward his preferred result, but at a certain point the entire argument structure begins to appear like a moral Rube Goldberg machine. I do not know what the structure of moral reality looks like, but it is unlikely to be a Rube Goldberg.<sup>181</sup>

### C. *Stephen Perry and a Different Form of Blameworthiness*

Professor Stephen Perry concedes that an agent loses control over the effects of action after completing the action, and that what happens next is a matter of luck, but denies that this truth is the best foundation for considering whether results increase blameworthiness.<sup>182</sup> We agree that the agent who causes the result is outcome-responsible for the result because the agent could have foreseen and taken steps to avoid the result by forbearing from the dangerous action. The question is whether the same arguments that show an agent is outcome-responsible and should therefore compensate a harmed victim in tort should also lead us to conclude that there is enhanced blameworthiness that justifies enhanced criminal punishment when the outcome occurs.

Perry's argument is this:

Given the way things turned out I have an extra increment of responsibility. . . . This is, to be sure, a different form of responsibility

---

181. It will come as no surprise that Katz's response to this analysis was to present me with three new variations on the hypotheticals. As the comedian Jerry Lewis used to say: “Alright already.”

182. Perry, *supra* note 94, at 94–95.



from blameworthiness. However, since I not only could have taken steps to avoid the outcome but, given the nature of my action, should have done so, it seems reasonable to think that in such cases outcome-responsibility has the effect of increasing blameworthiness. To put the point in the language of action, an act of deliberately killing another is inherently more blameworthy than an act of shooting at another with intent to kill.<sup>183</sup>

In criminal law cases, virtually all agents have the capacity to foresee that actions can cause prohibited harms and they always have the capacity to forebear. Thus, virtually all agents are action responsible. By precisely the same act of forbearance—fearing from performing the causally necessary last act—agents can prevent themselves from last-act attempting, which is punished despite the lack of an outcome, and from causing the result. Outcome-responsibility is not necessary to criminal law liability. In contrast, outcome-responsibility in tort is the very foundation of liability precisely because the aim of tort is to determine where the loss must fall. Perry does not really explain why action-responsibility is not sufficient for *criminal* desert, and compensation is anyway not the issue.

The heart of the argument is the last sentence, but the argument depends on a controversial act description. For the reasons given above,<sup>184</sup> I believe that the “act” of deliberately killing another is the same as the “act” of engaging in killing conduct with the intent to kill under conditions in which death will ordinarily occur. Surely they are morally the same. Thus, the only difference is outcome, not the act’s moral status. Outcome-responsibility may be crucial to tort liability, but it is not relevant to criminal liability based on desert.

#### D. *Other Arguments for Why Results Contribute to Desert*

This section briefly discusses two further arguments for why results contribute to desert: unjust enrichment and incompetence. I conclude that both arguments are plausible, but that neither is relevant to desert.

The unjust enrichment argument is psychologically straightforward but morally and legally problematic. The agent who succeeds obtains both the intended goal and the psychological satisfaction attending success; the agent who fails obtains neither.<sup>185</sup> Therefore, allegedly, the former has gained more and should receive more blame and punishment to right the moral ledger. To begin, the imposition of blame and punishment is not a remedy. Criminal punishment does not make the victim or society whole, nor does it give back anything that was improperly obtained. More important, all attempters are trying to succeed and are thus

---

183. *Id.* at 95.

184. *See supra* notes 182–83 and accompanying text.

185. H. L. A. HART, PUNISHMENT AND RESPONSIBILITY 131 (1968).

potentially guided *ex ante* by the moral and legal rules concerning success. No agent engages in criminal conduct with the hope or even the expectation of failure. The agent's knowledge that lesser punishment will be imposed if the agent luckily (for the victim) fails does not make the agent less culpable when the agent tries to succeed. The agent who fails has not intentionally "punished" himself or herself by failing. Agents who fail may feel less satisfied *ex post* than those who succeed, but such feelings do not affect the agent's culpability at the time of the criminal conduct. Moreover, if postcrime satisfaction is a genuine moral desert criterion, it should be applied to all criminals to calibrate more precisely the proportionate deserved punishment.

The incompetence argument suggests that agents who attempt and fail are in general likely to be less competent criminals than those who succeed. But except in cases in which the incompetence is sufficiently substantial to raise questions about the agent's normative capacity, I do not understand why competence has any bearing on desert. And once again, if the degree of competence bears on desert, it should be considered in all cases. A less competent agent who somehow succeeds might then be punished less harshly than a highly competent agent who somehow fails. But this seems a perverse result that should create even more doubt about whether competence is relevant to desert.

#### *E. Psychological or Sociological Necessity*

All common law and continental codes distinguish last-act attempts from completed crimes. And, empirical evidence strongly confirms that the belief in the importance of the distinction between attempted and completed crimes is robust among ordinary citizens.<sup>186</sup> One should exercise great caution before concluding that a universal legal rule and an apparently universal moral perception are rationally indefensible. Nonetheless, I have argued that it is not rational to distinguish last-act attempts on the basis of the agent's desert. Sanford Kadish agrees, but has argued that the rule may be so ingrained in our moral perceptions and attitudes that it would be useless to try to change the rule, and that doing so would create a loss of respect for the law.<sup>187</sup> If the distinction is so ingrained that trying to change the rule would have pernicious effects, then there are sound consequential reasons to retain the status quo. But I wonder if this is so.

Treating only last-act attempts, and not preparatory attempts, the same as successes would substantially soften the reluctance to adopt the scheme I propose. For example, careful examination of Robinson and Darley's empirical investigation of the attempt/completion distinction reveals that their survey did not include a last-act scenario, but did include

---

186. See ROBINSON & DARLEY, *supra* note 43, at 13–28.

187. Kadish, *supra* note 41, at 699–700.

a dangerous proximity test. If one consults the respondents' assessments of liability for completed offenses and various attempt scenarios, the liability scores for the dangerous proximity scenario were quite close to the completed offense, and both were substantially higher than for any of the other preparatory attempt scenarios.<sup>188</sup> Dangerous proximity tests will in fact criminalize conduct that is still quite far from the last act, but the dangerous proximity liability assessment that Robinson and Darley report is nonetheless close to that of the completed offense. Consequently, it is a plausible speculation that if the investigators had included a compelling last act scenario, the liability assessments might have been almost the same or even identical. In the risk creation study, holding the amount of injury risked or accomplished constant, there was almost no difference in liability assessments between cases in which there was a high probability of the risk occurring and those in which it did occur.<sup>189</sup> On the other hand, Robinson and Darley's study of "deviant" causal chains suggests that as the harm becomes more "remote," the subjects' assessment of criminal liability for the result approaches their assessment of liability for an attempt.<sup>190</sup> This study did not have a compelling last-act attempt as a control condition, however, and I wonder if a compelling last-act attempt scenario might not have produced *greater* liability assessment than a very remote success scenario.

Successes generally produce more resentment than last-act attempts, but it is not clear that they produce substantially greater liability assessments. If last-act attempts and successes are morally indistinguishable, and if the moral perception of the difference among ordinary citizens is not large—and anyway a sociological observation—perhaps the law could lead by abolishing the difference without compromising its own legitimacy.

## VI. CONCLUSION: CONSISTENT SUBJECTIVISM

Many commentators have recognized that the increasingly subjectivist view of criminal liability is inconsistent with the criminal law's emphasis on harm caused and its inclusion of result elements in the definition of some crimes.<sup>191</sup> The approach I am advocating eliminates such inconsistencies. Crimes would no longer have result elements, and all offenses would be defined only by intentional action and circumstances. There would be no difference in blame and punishment between conduct that causes harm and similar conduct that does not. Justifications would be assessed according to the reasonableness of the agent's subjective beliefs and the rightness of her reasons for action. Although such a crimi-

---

188. ROBINSON & DARLEY, *supra* note 43, at 19–27. The scale is logarithmic so the difference is not as close as it might appear.

189. *Id.* at 28–33.

190. *Id.* at 181–89.

191. *See, e.g.*, ROBINSON, *supra* note 129, at 109–11.

No. 2] REASON AND RESULTS IN THE CRIMINAL LAW 429

nal law would be almost entirely subjectivist concerning culpability, there would be no problem with legality because fully culpable action would always be required to justify criminal liability. The criminal law would blame and punish only those agents whose intentional actions risked a prohibited harm by violating a conduct rule established to prevent such risks and consequent harms. I recognize that this is a romantic vision of a coherent, consistent, and fair criminal law, but that is simply another virtue of the account.

## VII. APPENDIX: THE METAPHYSICAL CHALLENGE TO RESPONSIBILITY

In a causal universe, all phenomena, including intentional actions, are fully caused by the sufficient causes that produce them. In such a universe, why should the capacity to be guided by reason, or any other criterion, be the touchstone of responsibility? It is famously the case that many think that genuine responsibility is not possible at all.<sup>192</sup> This appendix considers this question by examining the metaphysical challenge to responsibility that was raised by the theme of luck that has been addressed throughout the article.

No analysis of this problem, or even statement of the question, could conceivably persuade everyone. There are no decisive, analytically incontrovertible arguments to resolve the metaphysical questions. Indeed, even theorists who adopt the same general approach to the metaphysical challenge substantially disagree.<sup>193</sup> Because one's position on this issue is metaphysically fundamental, however, I try in this appendix to sketch my approach. I argue that although there is no uncontroversial resolution to these issues, it is possible to provide a metaphysical, compatibilist account that positively explains our practices and normatively grounds them.

People are capable of practical reason and intentional actions, but they are also part of a universe that is massively lawful and regular, even if not fully deterministic.<sup>194</sup> Consequently, in principle, human action, including practical reason, can also be explained by mechanistic causes. Human action may be descriptively distinguishable from other phenomena, but perhaps the existence of practical reason makes no moral difference. Perhaps genuine moral responsibility is impossible in a largely mechanistic universe. After all, winds and tides are different phenomena, but they are both equally parts of the causal universe and are morally indistinguishable. If human action, like all other phenomena, is caused by antecedent events and by the causal laws of the universe, then perhaps people have no genuine freedom—no real choices or alternatives—even if they do act for reasons. If so, ascribing moral responsibility may be unjustified, and criminal responsibility might also be in doubt for those who believe, as virtually all do, that genuine moral desert is at least a necessary precondition for criminal blame and punishment. According to this view, human action, like everything else, is just part of the physical flotsam and jetsam of the universe—just a product of mechanis-

---

192. See, e.g., Strawson, *Impossibility of Moral Responsibility*, *supra* note 54, at 9–13.

193. See, e.g., Ishtiyaque Haji, *Compatibilist Views of Freedom and Responsibility*, in *THE OXFORD HANDBOOK OF FREE WILL*, *supra* note 23, at 203; Paul Russell, *Pessimists, Polyannas, and the New Compatibilism*, in *THE OXFORD HANDBOOK OF FREE WILL*, *supra* note 23, at 249–52.

194. Strawson refers to the recognition that the universe is massively lawful and regular as the “realism constraint” on theorizing about freedom. See Strawson, *supra* note 52, at 12.

tic biophysical forces—and thus furnishes no justifiable basis for ascriptions of responsibility.

Questions about why human action is morally unique have vexed philosophers for millennia, and the view one adopts has profound consequences for moral and legal theory and practice. The first standard answer to the claim that the universe is fully caused, “incompatibilism,” claims that mechanism is inconsistent with responsibility. Incompatibilism comes in two forms: hard determinism and libertarianism. The former admits that mechanism is true and therefore claims that no one is responsible for anything. The latter denies that mechanism is true as an explanation of human action, and therefore claims that responsibility is possible. The second standard answer, “compatibilism,” also embraces mechanism, but claims that responsibility is nonetheless possible. I will consider the two answers in order. In brief, I suggest that hard determinism can neither explain our practices nor ground a theory of desert, and that libertarianism is metaphysically implausible and also cannot explain our law. I argue that only compatibilism can explain and justify our moral and legal responsibility doctrines and practices, including the uniqueness of human action.

#### A. *Incompatibilism*

Let us begin with incompatibilism. Hard determinists and libertarians agree that “real” or “ultimate” responsibility is possible only if human beings have contra-causal freedom: the freedom to originate action not caused by prior events and influences. Otherwise, both agree, human beings have no real freedom, even if reasons are undeniably part of the causal chain that leads to action. They disagree, of course, about whether we have contra-causal freedom. Hard determinists believe that determinism or mechanism is true, that we lack contra-causal power, that acting for reasons is indistinguishable from other causes, and that therefore we are not responsible. Indeed, some hard determinists—the eliminative materialists—believe that the whole explanatory apparatus of folk psychology is false and treat mental states reductively as epiphenomenal byproducts of the biophysical firings of the brain and nervous system.<sup>195</sup> And if that is all that human action really amounts to, responsibility seems entirely unjustified. Libertarians believe that human beings are unique because we do have contra-causal power, and that acting for reasons originates uncaused, new causal chains. This godlike power of agent origination is the libertarian foundation of responsibility for action.

---

195. See CHURCHLAND, *supra* note 17, at 43.

### 1. *Hard Determinism*

Hard determinism generates an external, rather than an internal, critique of responsibility. That is, hard determinism does not try either to explain or to justify our responsibility concepts and practices. It simply assumes that genuine responsibility is metaphysically unjustified. Even if an internally coherent account of responsibility and related practices can be given, it will be based on an illusion.<sup>196</sup> To see why hard determinism or mechanism cannot explain our responsibility attributions, remember that such causal accounts “go all the way down:” determinism or mechanism applies to all people and all events. Thus, if determinism is true and is genuinely inconsistent with responsibility, no one can ever be really responsible for anything, and responsibility attributions cannot properly justify further action. But western theories of morality and the law do hold some people responsible and excuse others, and when we do excuse, it is not because there has been a little local determinism at work. For example, young children are not considered fully responsible because they are incapable of recognizing and properly weighing the right reasons for action and forbearance, not because they are determined creatures. Determinism does not loosen its grip on us as we age.

Determinism also does not tell us how we should act or what justice demands. Understanding the lawful regularities of human behavior might reveal what is possible for human beings and what is not, but it cannot dictate what morals, politics, and laws to adopt. As we know from history and ordinary observation, people are capable of adopting and acting on extraordinarily diverse moral, political, and legal schemes, which can make immense differences in the material and moral conditions of peoples’ lives. It is difficult to imagine what the metaphysical truth of determinism could tell us about how we should live. To what metaphysical facts should morality and law have to answer? Even if there are right answers, without placing ourselves outside the universe—anyway an impossibility—how would we ever know that we got it right?

“As if” responsibility is a classic hard determinist response to determinism’s inability to explain our concepts and practices or to justify alternatives.<sup>197</sup> “As if” admits that no one is really responsible and that no one really deserves anything as a matter of justice, but argues that responsibility is justified because we appear so committed to it and because responsibility practices seem to have the consequential virtues of encouraging good behavior and discouraging bad. As a practical matter, then, we should simply continue responsibility practices, including criminal blame and punishment, based on our knowledge of how people react

---

196. See SMILANSKY, *supra* note 56, at 40–73, 145–219 (arguing that free will is an illusion, but an illusion that is indispensable).

197. See, e.g., DANIEL C. DENNETT, *ELBOW ROOM: THE VARIETIES OF FREE WILL WORTH WANTING* 156–65 (1984); HERBERT PACKER, *THE LIMITS OF THE CRIMINAL SANCTION* 74–75 (1968).

to incentives such as the legal consequences of criminal conduct. There are two problems with this. First, it will not work if people believe that no one is really responsible. Only a genuine notion of responsibility will support the morality of giving people what they deserve. Few can be entirely cold calculators. Second, "as if" responsibility violates our most firmly held general principles of fairness and justice and would lead to morally unacceptable results in specific cases. How could we fairly praise and reward, or blame and punish, agents based on ascriptions of responsibility if responsibility is a myth? Specifically, a purely consequential view of punishment would not justify punishing anyone, no matter how heinous the deed, if we thought the overall consequences of doing so would be negative. More problematically, punishing the innocent to achieve maximum social utility would be justified in an "as if" moral universe. But core notions of justice would be violated by intentionally punishing an innocent agent, even if the public good would be promoted. These are familiar deontological objections to any purely consequential morality, but they have practical bite. Unless there is an extraordinary shift in our consciousness and attitudes, it will be difficult for us self-consciously to believe that we are acting fairly if attributions of responsibility are supported only by a mythical foundation.

Hard determinism can produce an internally coherent, forward-looking consequential system that treats human action specially and that might possibly encourage good behavior and discourage bad. Nevertheless, hard determinism cannot explain or justify our present blame and punishment practices, which are essentially retrospectively evaluative. What is more, hard determinism does not justify a morally acceptable system of just punishment because it is willing to impose pain based on a fiction, and because it leads to morally unappealing results in specific cases.<sup>198</sup> In sum, hard determinism cannot explain or justify the relation between action and desert.

## 2. *Libertarianism*

The libertarian concedes that determinism or mechanism may account for most of the moving parts of the universe, but argues that human action is not subject to causal influence. The libertarian is thus able to distinguish action from all other phenomena and to ground responsibility in contra-causal freedom. If the libertarian is correct, then the buck really does stop with human intentions, at least if intentional action is both rational and not influenced by coercive threats. The difficulty is that libertarianism produces a worthless view of responsibility and depends on a "panicky" metaphysics, to use P. F. Strawson's phrase.<sup>199</sup>

---

198. See WALLACE, *supra* note 57, at 54–61.

199. Strawson, *supra* note 59, at 80.



Moreover, the criminal law is often inconsistent with a libertarian metaphysics of action.

One form of libertarianism holds that human actions are the product of indeterministic events in the brain. But why should such events ground responsibility? In what way are actions produced by indeterminate or random brain events “ours,” or an exercise of a freedom worth wanting? If our brain states and the world in general are truly random or indeterministic, then it is difficult to imagine how responsibility could be possible. Such randomness is deeply implausible, however. We could not explain the regularity of physical events and of human interactions, nor could we explain the dependable relation between intentions and actions, unless there was a cosmic coincidence of astounding proportion that accounted for the regularity. More important, brains would be akin to random number generators, and our behavior would be equivalent to the random numbers generated. This is scarcely a secure foundation for responsibility or for any form of moral evaluation, because rational intentions would be random rather than the product of genuine practical reason. In a sense, nothing would be “up to us.” There would be essentially no relation between our intentions, our actions, and their outcomes. Moreover, if our brain states and the world in general are random, we should not believe that our lives will go any better for us. Unless the matter in our brains and bodies rearranges itself, we will not be any different from the way we are. A random world would be impossible to manipulate and control dependably. In sum, if rational action is simply a product of biophysical indeterminacy, no one should be responsible for any action.

An apparently more plausible version of libertarianism concedes that prior events and experiences affect our mental states, but alleges that our actions are ultimately not caused by anything other than ourselves. To begin, such a theory cannot specify the nature of the causal relation between an agent and an act that she causes.<sup>200</sup> Furthermore, there is no observational evidence that would confirm that agent-causation is true. Our experience of psychological freedom when we act intentionally is often cited, but such arguments from experience for contra-causal freedom have been properly criticized on many grounds.<sup>201</sup> Perhaps most important, our experience might simply be a psychological fact about us rather than good evidence to justify the truth of agent-origination. Moreover, why would we choose to adopt such an implausible theory when there are also no good nonobservational grounds, such as coherence or parsimony, for doing so?<sup>202</sup> Finally, our desire to believe

---

200. See Michael S. Moore, *The Metaphysics of Causal Intervention*, 88 CAL. L. REV. 879, 884 (2000).

201. See, e.g., John Stuart Mill, *An Examination of Sir William Hamilton's Philosophy*, in J. M. ROBSON, 9 COLLECTED WORKS OF JOHN STUART MILL 449-57 (1979).

202. BOK, *supra* note 5, at 45.

that agent-causation is true, or its importance to our self-conception, is not independently a good reason to accept it as truth if the metaphysical foundation is false.

Even if agent-causation were true, the problems for responsibility that mechanism produces may remain.<sup>203</sup> Suppose that mechanistic forces determine all our mental states and processes other than our intentions themselves and suppose further that all these mental states and processes lead inevitably to the intentions that we form. Even if those intentions were somehow uncaused, this form of agent-causation would be both implausible and essentially indistinguishable from mechanism. Nor would responsibility be more secure if agents somehow always formed intentions not to do what they had reason to do based on prior deliberation, or if they acted for no reason at all despite having good reasons for some course of action. Why would the freedom always to choose a course of action that, all things considered, we thought we should reject, or the freedom always to choose to act for no reason, be a freedom worth wanting? Why would we want to act intentionally for no good reason? So, if we always acted in accord with the reasons we were caused to have, the substance of our actions would be determined, even if the ultimate intention to act was somehow uncaused. And if we acted for no good reason that we credited, this freedom of action would be undesirable because it would not involve independent, rational evaluation of what we had reason to do.

Finally, suppose a libertarian concedes, as one rationally must, that preceding events and states affect our mental states and processes and thus also our reasons for action. In that case, external causal processes would not be the cause of action only if some aspect of the agent evaluates the person's reasons for action independently of the reasons the external world causes the agent to have.<sup>204</sup> But then the agent would not be aware of the aspect of oneself that is performing the independent evaluation. If causal forces are causing an agent to perceive the world a certain way and to have certain reasons for action, then the agent will not question that perception or those reasons because there would be no reason to do so. An independent, evaluating aspect of the self is not identical to the conscious, intentional agent because that agent would not be aware of any reason to evaluate its reasons independently of those it was already determined to have.

Suppose there is an aspect *of the agent* that is able to cause the agent to perceive the world differently and to have reasons for action different from those produced by external causal processes. According to the libertarian account, that aspect of the agent must then be able to override the reasons produced by external causes that the agent consciously has. But in that case an agent's reasons for action are being con-

---

203. *Id.* at 48–50.

204. I owe this argument to Hilary Bok. *Id.*

trolled by an aspect of the agent of which the agent is unaware, and with which the conscious, intentional agent does not identify. This nonconscious aspect of the agent is thus controlling the agent's reasons for action as much as external forces ever did. The same threat to moral responsibility therefore obtains as if mechanism were entirely true. And, finally, we are entitled to ask the libertarian why an independently evaluating aspect of the self is uninfluenced by preceding events and how it comes to have contra-causal power. In sum, even if our intentions are not determined by external causal processes, libertarians must concede that they are nonetheless causally determined by processes over which conscious agents have no control. Consequently, in this version of libertarianism, too, human action is the product of causal forces and is indistinguishable from the other phenomena of the universe.

No one, alas, is a prime mover unmoved—a being with contra-causal freedom that can rival or override the world of causes in which we live. Metaphysically, it is preposterous to suggest that human action by rational adults is uncaused and that we are thus responsible for it, whereas we are not responsible for any other phenomena of the universe—including the results of action—because they are the products of mechanistic, causal laws. Many practical lawyers find such metaphysical claims unsettling or irrelevant and simply suggest that, whatever the metaphysical truth may be, the criminal law assumes that rational human action is free and no more needs to be said. But much as we are right to reject the “as if” responsibility that hard determinism offers because it is a fiction, we should also reject as rationally indefensible the equal fiction of libertarian freedom as the basis of criminal responsibility.

Moreover, the criminal law is but weakly and inconsistently dependent on libertarian assumptions. A main exception, of course, is accomplice liability, which obtains even if the potential accomplice's aiding or abetting behavior did not causally contribute at all to the perpetrator's offense.<sup>205</sup> If the actual perpetrator's intention to offend is entirely uncaused by external variables and thus entirely up to the perpetrator, why should a person who encourages or aids be guilty derivatively, even if the accomplice does help? After all, the accomplice did not in any way cause the perpetrator to commit the crime, because only the perpetrator caused himself or herself to commit the crime. We might wish to criminalize aiding and abetting behavior as an independent crime, especially if the aid furnished was necessary for the perpetrator's success. But guilt would no longer be derivative of the perpetrator's culpability. We punish accomplices derivatively, however, because we do believe that an ac-

---

205. The analysis of accomplice liability's inconsistency with libertarianism applies a fortiori to inchoate crimes such as solicitation and conspiracy. Attempt presents a special case because I have claimed that many important cases of attempt do not represent inchoate crime. See *supra* Part IV.A (considering attempt doctrine in detail and distinguishing “last-act” attempts from “preparatory” attempts).

complice's behavior does or potentially does causally contribute, thus undermining a fully libertarian basis for criminal liability.<sup>206</sup>

Libertarianism also cannot explain the distinction between culpable and excused conduct. For example, there is no reason to believe the conduct of people with severe mental disorders or the conduct of those acting under duress is any less subject to contra-causal freedom than the conduct of fully rational agents or those not acting under duress. Even if the conduct of the mentally ill is influenced by, say, abnormal neurotransmitters, why should this affect contra-causal freedom? After all, the behavior of sane people is influenced by normal neurotransmitters. No one has answers to such questions, but the implausibility of neurotransmitters being the source of contra-causal freedom is as great as the implausibility of libertarianism itself. Mentally ill and sane people alike are practical reasoners who form and act on intentions based on their desires and beliefs. We might excuse some mentally ill people because they exercise their contra-causal freedom to act for irrational reasons, but in that case rationality and not contra-causal freedom is the criterion that determines culpability.

One might suggest that the criminal law need not be consistent about its philosophical presuppositions, but we cannot have justifiable moral confidence in a criminal justice system that is inconsistent in its metaphysical justifications for blame and punishment. Moreover, as we shall see presently, compatibilism is consistent with all criminal law doctrines. In sum, libertarianism, unlike hard determinism, is metaphysically implausible, but like hard determinism, it cannot explain our doctrines and practices.

### B. *Compatibilism*

The compatibilist answer to the argument that mechanism and moral responsibility are inconsistent is to claim that mechanism is both true and compatible with moral responsibility. This answer in its various forms is widespread among moral philosophers<sup>207</sup> and is the approach that I and many other criminal lawyers explicitly or implicitly adopt.<sup>208</sup> This approach accepts completely that we live in a thoroughly causal world, at least at the macro level, and that causal processes produce human action and all the other phenomena of the universe. But it also holds that genuine responsibility is possible. This approach best explains

---

206. See *supra* Part IV.D (discussing accomplice liability).

207. See, e.g., BOK, *supra* note 5, at 75–91, 129–31, 146–51; JOHN MARTIN FISCHER & MARK RAVIZZA, S.J., *RESPONSIBILITY AND CONTROL: A THEORY OF MORAL RESPONSIBILITY* 207–39 (1998); WALLACE, *supra* note 57, at 7–16; Haji, *supra* note 193, at 202–03; Russell, *supra* note 193, at 229. See generally John Martin Fischer, *Recent Work on Moral Responsibility*, 110 *ETHICS* 93 (1999) (surveying recent work among moral philosophers).

208. See generally Michael S. Moore, *Causation and the Excuses*, 73 *CAL. L. REV.* 1091, 1121 (1985); Morse, *supra* note 28, at 337–62.

and justifies our moral and criminal law practices without endorsing the implausibilities of libertarianism. Even if mechanism is true, the law's concepts of moral responsibility and deserved blame and punishment are rationally defensible in the compatibilist view.

Defending this position fully goes far beyond the goals of this article, and there are many different defenses of compatibilism.<sup>209</sup> Because compatibilism is the source of this article's normativity, however, I shall sketch a defense.<sup>210</sup> Before doing so, though, I must concede again that currently (and perhaps forever) a definitive answer to whether responsibility is consistent with determinism or mechanism is impossible.<sup>211</sup> The most we can do is try to justify our concepts and practices by using the best moral and political theories available and by trying to be as faithful as possible to the facts about the universe, including human behavior, as we know them. No account, however, will ever cause determinist anxieties to disappear or will ever provide universally satisfying, uncontroversially persuasive answers. Because I accept the realism constraint, the most my defense can do is to offer an internal account of responsibility that is not inconsistent with the truth of determinism or of accepted principles of fairness. Thus, in what follows, I will try to show that a coherent, internal account of responsibility can be given that depends on few controversial assumptions about human beings, that explains our current practices, and that accords with deeply held principles of fairness.<sup>212</sup> Nevertheless, at some point the arguments run out and there is dialectical stalemate. At that point, the consistency of compatibilism with our understanding of the facts and our legal and moral practices becomes a strong argument in its favor.<sup>213</sup>

My defense of compatibilism begins with the unprovable but common and plausible assumption that morality and its practices, including blame and punishment, are human constructs and that there are no

---

209. See sources cited *supra* notes 207–08 (providing examples of different defenses).

210. The defense has been most influenced by the views of P. F. Strawson, Jay Wallace, and Hilary Bok, although none of them would fully accept my account.

211. See NAGEL, *supra* note 50, at 112. Indeed, answering all hard determinist criticisms would require “[s]ubstantive metaphysical and modal discussion of counterfactuals, causation, and laws of nature.” Kadri Vihvelin, *Stop Me Before I Kill Again*, 75 PHIL. STUD. 115, 142 (1994) (claiming that compatibilism is true). This is a task far beyond the scope of this article and the criminal law's needs. Even if it were accomplished, I freely concede that it would not convince entrenched hard determinists.

212. This form of justification of compatibilism has been characterized as compatibilism motivated by “praxis,” because it treats a rational legal system as a moral good, and as motivated by “underpinnings,” because it provides an internal account. See RICHARD DOUBLE, *METAPHILOSOPHY AND FREE WILL* 28–29 (1996).

213. See Charles Larmore, *Beyond Religion and Enlightenment*, 30 SAN DIEGO L. REV. 799, 812–14 (1993) (arguing that existing moral convictions require revision only if such convictions appear inconsistent with other moral convictions or with nonmoral information that seems fixed). At such junctions, Wittgenstein famously said: “If I have exhausted the justifications I have reached bedrock and my spade is turned. Then I am inclined to say: ‘This is simply what I do.’” LUDWIG WITTGENSTEIN, *PHILOSOPHICAL INVESTIGATIONS* § 217 (G. E. M. Anscombe trans., 3d ed. 1958) (commenting on how one justifies obeying a rule).

metaphysical moral facts about the universe to which our moral judgments and practices may appeal. For example, some claim that strong, contra-causal freedom is really necessary for responsibility,<sup>214</sup> but I have no idea how one would ever metaphysically confirm the truth of such a proposition. Although the assumption that morality is a social construct rejects extreme metaphysical moral realism, it does not commit one to accepting relativity and subjectivity in morals. There are better and worse arguments for almost any position, and to abandon the quest for greater rationality and agreement in morals is to reduce moral judgment and practices to the unalloyed expression of emotion and power. I also believe that incredibly refined, hypertechnical arguments using modal logic to analyze the meaning of phrases such as, “could have done otherwise,” when used to prove that agents could have acted differently in a causal world, are simply unable to bear the moral weight they are asked to support.

The next step in the defense of compatibilism draws the common distinction between theoretical and practical reason and considers the implications of the position that practical reason is inevitable. Theoretical reason is concerned with how the world really is: the apparently scientific world view. The latter is concerned with what human agents have reason to do. For example, whether human beings in fact deliberate and act for reasons is a question of theoretical reason. Do they or don't they? Whether it *should* matter to us if they do is a question of practical reason.

It is simply a fact about human beings that they do deliberate, they do act for reasons, and, even if our actions are determined, they make a difference. Actions can cause pleasure or pain, can create wealth or poverty, can be kind or cruel. We discuss, argue, fight, and even kill over morals and politics precisely because moral and political regimes make an enormous difference to our well-being. And it is unimaginable that conscious, social, deliberative, and intentional creatures such as ourselves could live in a world without moral and nonmoral norms, without some moral and political regime. Even if determinism is true, we cannot wait for it to happen. We must determine what determinism dictates.<sup>215</sup>

---

214. See Strawson, *Impossibility of Moral Responsibility*, *supra* note 54, at 5–7.

215. Stephen J. Morse, *Waiting for Determinism to Happen* 5 (1999) (previously available at <http://www.legalesays.com>; now on file with author).

The notion that we cannot wait for determinism to happen, that we cannot slavishly and unthinkingly follow our fate, is beautifully expressed in Monica Ali's acclaimed novel, *Brick Lane*. The protagonist, Nazneen, is an eighteen-year-old village “girl” from Bangladesh whose father hastily arranged her marriage after her younger sister, Hasina, ran off and made a love match. Nazneen's husband, Chanu, is a forty-year-old Bangladeshi who took her immediately to London. In the following passage, Nazneen is thinking to herself as she ponders her new life in a poor Bangladeshi neighborhood in London:

It worried her that Hasina kicked against fate. No good could come of it. Not a single person could say so. But then, if you really looked into it, thought about it more deeply, how could you be sure that Hasina was not simply following her fate? If fate cannot be changed, no matter how you struggle against it, then perhaps Hasina was fated to run away with Malek. Maybe she strug-

There is no possibility that we could have perfect knowledge of future behavior without changing that behavior. There can be no determinist oracle we can consult to eliminate the need to deliberate.<sup>216</sup> Only creatures who did not care about their lives at all or who believed that their deliberations and intentions made no material difference would fail to deliberate, but precious few agents fit this description. It is almost unthinkable for virtually anyone to live a life without any purpose whatsoever. As long as we do have purposes or ends, we have reason to ask what we have reason to do. Within the inescapable realm of practical reason, deliberation and intentional action are inevitable.

The next crucial compatibilist assumption is that most people have the full, general capacity to grasp and be guided by good reasons by the time they reach the "age of reason." By a general capacity, I mean simply the underlying ability to engage in certain behavior. For example, English speakers have the general capacity to speak English, even when they are silent. General capacities have ranges and limits of course, but within the range, we believe that people are capable of acting in a particular way. Hard determinists might argue that the only "capacity" we genuinely have is the momentary capacity to do what we are in fact doing at the moment. From the vantage point of theoretical reason, this may be right, but no one knows. More important, to reject the notion of a general capacity would do violence to a fundamental, common sense understanding of human behavior that guides interpersonal interaction and that is central to moral evaluation. The general capacity to grasp and be guided by good reason—normative capacity—does not mean that all of us behave rationally all or even most of the time. People often act non-rationally, irrationally, foolishly, and the like. But like the silent English speaker, we retain the capacity to be guided by reason even if we are not exercising that capacity.

Because the capacity to be guided by reason or rationality is fundamental to my account of compatibilist responsibility, one might desire a precise, uncontroversial definition of rationality. But such a desire would be unreasonable. There is no uncontroversial definition in philosophy or the social sciences. Nonetheless, the implicit common-sense notion is that the capacity for rationality is a congeries of skills, including the ability to perceive accurately, to reason instrumentally according to a minimally coherent preference-ordering, and to appreciate the significance of reasons and their connection to our actions. This notion of rationality is

---

gled against *that*, and *that* was what she could not alter. Oh, you think it would be simple, having made the decision long, long ago, to be at the beck and call of fate, but how to know which way it is calling you? And there was each and every day to be got through. If Chanu came home this evening and found the place untidy and the spices not even ground, could she put her hands like so and say, "Don't ask me why nothing is prepared, it was not I who decided it, it was fate." A wife could reasonably be beaten for a lesser offense.

MONICA ALI, BRICK LANE 9 (2003).

216. See BOK, *supra* note 5, at 79–88.

central to and grounded in our ordinary, everyday understanding of practical reason and its critical role in human interaction, including moral evaluation and practices. Our understanding of rationality is always open to revision, but we all use a common-sense understanding of it to evaluate the nonmoral and moral conduct of ourselves and others. It could not be otherwise among consciously deliberating creatures. If one wishes to abandon the role of rationality in human conduct, then the burden should be on those who wish to reject it to provide a compelling account to do so.

Neither the concept of a general capacity nor the specific capacity to be guided by reason is inconsistent with the truth of determinism. Indeed, such capacities are surely the product of mechanistic events, such as evolutionary natural selection and other biophysical processes. Furthermore, the truth of determinism does not mean that people are puppets or otherwise nonrational or nonintentional creatures. People and puppets both exist as the result of causal processes, but they are critically different from one another. No Geppetto pulls our strings. We may be secure in our assumption that human beings are intentional agents with the general capacity to be guided by reason.

A compatibilist holds an agent responsible when the agent acts intentionally by either complying with or breaching an accepted moral or legal obligation, and when the agent is generally capable of grasping and of being guided by reason in the context.<sup>217</sup> If these conditions are met, then praise and reward or blame and punishment are morally justified. For the compatibilist, moral responsibility is dependent primarily on the agent's general capacity to grasp and be guided by reason, because it is reasonable and fair to hold an agent responsible only if the agent possesses this capacity. This does not mean that every agent is always guided by good reason. All too often, even the best and most well-endowed agents are not guided by good reason. Rather, it means simply that we assume that the general capacity is present, even when it is not exercised or when, through no fault of the agent, circumstances for its exercise are not propitious. At the normative extreme we might decide that an agent's general capacity is so weak or compromised that we will not hold the agent responsible for failing to be guided by good reason, but this will be the exception, not the rule. After all, most agents do not suffer from impaired intelligence, major mental disorder, or other conditions that compromise the general capacity to be guided by reason. Most agents can be guided by good reason and are perfectly capable of understanding the relatively simple prohibitions at the core of criminal law.

A compatibilist excuses if the agent does not act or if an agent who acts lacks the general capacity for rationality in the context.<sup>218</sup> According

---

217. See WALLACE, *supra* note 57, at 86–87, 157–61.

218. A compatibilist account would also accept coercion or compulsion as an excusing condition, even if the agent retained the capacity for rationality and acted intentionally. See MODEL PENAL



to this account, human action is different from the rest of the causal universe not because it is uncaused, but rather because it is the product of potentially rational practical reason. Only people act for reasons and only action can be fully guided by reason.

The compatibilist account I have given of responsibility accords with our most firmly held and most persuasive theories of fairness.<sup>219</sup> Morality and law are fundamentally action-guiding, and thus are addressed only to creatures capable of practical reasoning. It is coherent to hold people responsible only if agents act and are capable of practical reasoning that can properly use morality and law as guides. We believe that it is fair to praise and blame, and to reward and punish, precisely when intentional actions that cause good and evil have been performed by agents capable of rationality. If an agent does not act, then the person's bodily movement cannot be attributed to him as an agent and thus it would be bootless to hold the agent morally responsible.<sup>220</sup> Only actions can be guided by reason. Moreover, if an agent is incapable of rationality (or is compelled to act), once again it would be unfair to ascribe responsibility because the agent either cannot be guided by good reason (or faces an unfairly hard choice). Although, as we have seen, a consequentialist might be willing to reward or punish in cases in which the agent lacks the general capacity for rationality or the agent faces an unfair choice, such a forward-looking scheme both fails to explain irreducibly deontological aspects of current morality and it famously entails further consequences—such as the potential deliberately to punish the innocent—that few would accept.

Our most deeply held principles of fairness hold that it is neither gratuitous to praise and reward appropriately for intentionally satisfying accepted moral and legal expectations, nor cruel to blame and punish appropriately for intentionally breaching such obligations. Furthermore, we currently adhere to no principle of fairness that is related to the truth or falsity of determinism. Some might wish to adopt such a new principle, but doing so would be based on an external critique of current morality.

Even if an internally coherent and apparently fair account of responsibility that is consistent with determinism is possible, perhaps it is all just an illusion. Perhaps it is true that our fates are sealed by the

---

CODE § 2.09(1) (1962) (excusing actors who are coerced by the threat of unlawful force that a person of reasonable firmness would be unable to resist). Such cases are relatively rare, however, and the justification for the excuse has nothing to do with causation or determinism. We excuse agents acting under duress, for example, because the agent is wrongfully placed in a woefully hard choice situation—"Do wrong or else!"—and we think it would be unfair to hold the agent responsible for acceding to the threat. SMILANSKY, *supra* note 56, at 87.

219. Cf. DWORKIN, *supra* note 49, at 285–303 (insisting on the importance of distinguishing the results of chance, for which we are not responsible, from the results of choice, for which we are responsible, because the distinction is crucial to our conception of morality and of ourselves).

220. I leave aside the complicated question of omissions, but they, too, can be accommodated within the theory presented.

causal laws of the universe, that we do not really deliberate, that our intentions and actions do not really affect the world. Such conclusions would erroneously confuse determinism with fatalism, which, roughly, is the doctrine that our intentions and actions do not make a difference. Our fate is not a separate entity that buffets us about; our intentions and actions *are* our fate.<sup>221</sup> Determinism or mechanism does not deliberate for us. We do have the general capacity to reason, often exercise that capacity, and thereby affect the world. For those with the most incompatibilist, hard-deterministic intuitions, this answer will seem insufficient. They will deny that “real responsibility” is possible if determinism or mechanism is true.<sup>222</sup> Again, however, responsibility attribution and related practices are moral enterprises, and it is difficult to imagine what metaphysical facts external to our normative world morality should answer to. I contend that responsibility based on our general capacity to reason and to act in accord with reason is the best explanation of current practices and is sufficiently “real” morally to justify the practices based on it.

Even if such an account is *internally* coherent, is it desirable? One could argue either that another account of responsibility is preferable or that we should dispense with responsibility altogether. I have suggested that the account offered best explains our practices and is consistent with profound and widely held moral principles of fairness. I shall leave to others the task of suggesting why a different account would better explain current practices or would be more desirable, but I do believe that some regime of responsibility is worth maintaining.

What makes us distinctively human is our capacity for reason, which in turn makes us both capable of genuine normative evaluation and appropriate objects of such evaluation. Evaluative moral responsibility is crucial to our sense of ourselves as persons, as agents, as objects of dignity and respect. And moral responsibility coheres with other moral notions of supreme importance, such as desert, justice, and fairness. Other sentient creatures can suffer and deserve to be treated without unnecessary pain, and some other creatures have highly developed intelligence and culture, but we are the only creatures capable of leading fully moral lives. Responsibility is central to such a life and contributes to a meaningful life of dignity and worth. Responsibility enriches our lives; abandoning it would diminish personhood.

---

221. See BOK, *supra* note 5, at 76–77.

222. See, e.g., Strawson, *supra* note 50, at 7. Strawson asks us to imagine if compatibilist responsibility would be sufficient to condemn a genuine evildoer, say, Hitler, to eternal torment in Hell. This thought experiment requires the theological view that God gives human beings the godlike power of agent causation. If so, as Dante famously thought, eternal torment might be considered fitting and deserved. See DANTE ALIGHIERI, PURGATORIO XVI: 67–83 (Jean Hollander & Robert Hollander trans., 2003). But would compatibilist responsibility be enough to justify such punishment? I think not, but for a reason unrelated to the comparative validity of compatibilist responsibility. My view, unlike Dante’s, is that nobody deserves such cruel punishment. Even if I were convinced that secular agent causation were possible, I would hold the same view.

Responsibility is desirable also because it contributes to the creation and maintenance of moral communities.<sup>223</sup> Agency, responsibility, and desert are all moral notions, inextricably intertwined in our moral and nonmoral lives and interactions. When we attribute responsibility and appropriately express those attributions through practices of praise and blame, reward and punishment, we affirm and deepen our commitment to common moral obligations that bind us together. To diminish or abandon a robust notion of responsibility would be to weaken those ties and the communities that nurture us.

The reasons that I have given for maintaining a system of responsibility are potentially persuasive only if responsibility is not a myth. The compatibilist has no greater warrant to adopt "as if" responsibility than the hard determinist. As this appendix has argued, however, responsibility based on the capacity for reason should be sufficiently real to satisfy all except those who believe that in a thoroughly causal world there is no genuine morality that matters. Nonetheless, I recognize that the hardest of the hard determinists will not be convinced and that the libertarians will insistently claim that they really do have contra-causal freedom.

Finally, compatibilism is fully consistent not only with deeply held moral principles; it is also fully consistent with criminal law. The criminal law's concept of a responsible person is precisely that of an intentional agent with the general capacity for rationality, and thus compatibilism is consistent with both doctrines of culpability and doctrines of excuse. On occasion, the criminal law, which is not entirely driven by desert justifications, decides on consequential policy grounds to dispense with or diminish culpability requirements. Classic examples are strict liability crimes, the rejection of excuse for nonculpable ignorance of the law, and felony murder liability. But these doctrines are exceptions to the general principle that there should be no blame and punishment either without desert or disproportionate to desert. Many find such doctrines morally problematic. When the law does adopt them, it does so not because it rejects compatibilism and endorses hard determinism, but because it accepts a thoroughly consequential justification for punishment. Compatibilism is fully consistent with a self-conscious, mixed justification for punishment that amalgamates retribution and consequentialism. Virtually always, however, desert remains as at least a necessary condition of just punishment, and compatibilism's emphasis on the general capacity for reason as the touchstone for responsibility furnishes the best understanding of why this should be.

---

223. SMILANSKY, *supra* note 56, at 88-89.

Individual empowerment, autonomy, and self-respect depend on the existence of a Community of Responsibility. Such a community requires recognition of the connection between inherent respect for a person and concern for his or her agency. . . . [C]ontrol compatibilism is the most adequate defence of our dominant understanding of the Community of Responsibility and the idea of criminal justice as formulated by H. L. A. Hart.

*Id.*