

## THE PROBLEMS WITH BLAMING

By Theodore Y. Blumoff<sup>#</sup>

“To . . . blame is not to judge merely that what happened was a . . . bad thing, but to judge the person for having done it, in view of the circumstances under which it was done. The difficulty is to explain how this is possible.”<sup>\*\*</sup>

[H]ow the reformatory view of punishment can be accused of disrespect for human personality, because . . . it uses a man’s animal organism or his lower psychological nature as a means to the good of his higher self, I cannot profess to understand.<sup>\*\*\*</sup>

### Introduction: On Blaming

We devote substantial time and energy to blaming wrongdoers, and too often (in the process) we brush aside our basic commitments to the decency of compassion that produces the goal of rehabilitation. Although we cannot escape the practice of blaming because there is something in our nature that requires the practice,<sup>1</sup> we must be able to harness it. I, like Hastings Rashdall who is quoted above, don’t profess to understand those who oppose reformatory punishment because it interferes with the goal of retribution. Protecting our children, in my view, transcends retribution *qua* retribution. I will advance a more pragmatic stance when it comes to protecting future generations, one which hopes to bring moral educational tools to those adjudicated wrongdoers who are capable of being helped.

I should be clear that I do not suggest that we eliminate blaming as a social practice. To the contrary: It endures as a feature capable of pervading our decision-making and so our culture because it permits us to tell wrongdoers, as directly as the context allows, that

---

<sup>#</sup> Professor of Law, Mercer University School of Law, Macon, Georgia. Portions of this paper were delivered at a colloquium at Georgia State University College of Law, at a Faculty Workshop at Washington University’s School of Law, at John Marshall Law School, and at a Mercer University Law School Faculty Development Colloquium. I will present the main thesis of this work at a conference on “Law, Mind and Brain,” at University College London and the Gruter Institute on February 13-14, 2006. Many thanks to those colleagues and friends who offered friendly and helpful comments and suggestions on this and related projects, including Robert Audi, Karen Kovach, Harold Lewis, Stephen Morse, and Jack Sammons. As always, I appreciate the financial and collegial support of Mercer Law School. Finally, special thanks to Rebecca McKelvey (Class of 2006) for timely and thorough research assistance.

<sup>\*\*</sup> Thomas Nagel, *THE VIEW FROM NOWHERE* 120 (Oxford 1986).

<sup>\*\*\*</sup> See Hastings Rashdall, *Punishment and the Individual*, in G. Ezorsky, ed., *PHILOSOPHICAL PERSPECTIVES ON PUNISHMENT* 6, 7 (New York: State University of N. Y. Press, 1972).

<sup>1</sup> On the naturalist features of blaming, see Adam Smith, *THE THEORY OF MORAL SENTIMENTS* (Prometheus Books 2000); Peter Strawson, *Freedom and Resentment*, *PROC. BRIT. ACADEMY* (1962), reprinted in *FREE WILL* 59 (ed. by Gary Watson; Oxford 1982).

they have committed wrongs that we disapprove of, that they should have done something else on the occasion we judge wrongful, and that they should not commit similar acts in the future (or else face substantial pain, if they are not facing it already). The need to cast blame is both a natural defensive stance and, I think, an essential one, at least for the foreseeable future. To the question, then, “Is blaming per se inappropriate,” where blaming instantiates an ascription of fault or demerit to an individual for producing an unwelcome outcome intentionally (at least usually),<sup>2</sup> the answer is: “Of course not. It can’t be.”

It can’t be, because if blaming served no other purpose, it would remain a valued tool in our efforts to advance our children’s ethical and moral conditioning. As a child-rearing practice exercised with love and compassion, and employed with a long term view of moral development, blaming often helps us achieve the important goal of educating our children in shared moral, social, cultural and legal norms, thereby furthering the project of civilizing them in light of our status as the ultimate primate.<sup>3</sup> Anyway you look at it, we are members of a species of the primate genus that needs civilizing care.

As good parents, we have generally devised and followed the basic formal outlines for the moral education of any child, and not only our own. It begins with a certain kind of discipline that requires, as a necessary condition to success, the neurobiological *capacity* to internalize the basic moral demands of our society, and thereafter some  $x+$  number of *opportunities* to learn the basics, where  $x+$  often represents a number of examples over a period of time that falls well outside our initial uneducated hopes and expectations, at least for some of our children some of the time. To meet our own needs and those of our children, we employ educational strategies that require perseverant efforts over long periods of time directed at teaching the everyday material of values. Reaching that goal, in turn, dictates the need to make available to every child variably graded exercises pointed at learning the nature of individual agency and its obligations: at instilling the notion that each of us can (prima facie) recognize ourselves as sources of effect on the world, and so at inculcating practices that appropriately praise good and censure bad.<sup>4</sup> An adequate moral

---

<sup>2</sup> I don’t mean to suggest that inadvertently caused harms are entirely non-blameworthy, see Judith Andre, *Nagel, Williams, and Moral Luck*, 43 ANALYSIS 202 (1983), only that they typically occasion less blame than do intentionally caused harms..

<sup>3</sup> In the child-rearing context, blame serves as a tool for teaching moral responsibility, including genuine feelings of shame and guilt and resentment and the like, on the one hand, and love and respect and compassion, on the other. Following Hume, therefore, I incline toward a utilitarian (even pragmatic) view of blame and punishment. See David Hume, *Of the Immortality of the Soul*, in SELECTED CLASSICS 324, 327-28 (Stephen Copley and Andrew Edgar eds. Oxford 1996). See generally Paul Russell, *Hume on Responsibility and Punishment*, 20 CAN. J. PHIL. 539, 550-52 (1990).

<sup>4</sup> This course follows multiple tracks, and not least ones that describe form and another substance. There are, as Bernard Williams notes, many modalities for blaming. Some take the proleptic form, “*If I were you I would have (or not have)  $\Phi$ ed.*” This form of advising involves “treating the person who is blamed like someone who had a reason to do the right thing but did not do it.” Bernard Williams, *Internal Reasons and the Obscurity of Blaming*, in MAKING SENSE OF HUMANITY AND OTHER PHILOSOPHICAL PAPERS 35, 42-3 (Cambridge 1995). On the substance side, the least that we can and should expect is a sufficient understanding, one that takes into account Hillel’s

education permits the student with sufficient capacity to internalize a substantial measure of what it means to *hold* oneself, and *be held*, responsible for chosen conduct.<sup>5</sup> We usually know (or at least believe we know) with respect to our own children that when they make morally relevant choices, they possess the fundamental capacity-in-fact to make important decisions with at least a reasonable measure of prudence if not yet genuine moral sensitivity.<sup>6</sup> This is a central concern for us as parents and bearers of our culture.

But blame attribution processes also operate when we find fault with individuals who are *not* our own wonderful children. We assign blame to these *others* when we are persuaded that they have engaged in misconduct that we condemn and criminalize. Attributing blameworthiness and criminal wrongdoing to someone whose lives we visit only briefly and in tragic circumstances – and whose childhoods and opportunities for moral

---

description of TORAH: “*What is hateful unto you, do not unto your neighbor. The rest is commentary—now go and study.*” See, e.g., Karen Armstrong, A HISTORY OF GOD: THE 4000-YEAR QUEST OF JUDAISM, CHRISTIANITY AND ISLAM 72 (Alfred A. Knopf 1993); Joseph Telushkin, JEWISH LITERACY: THE MOST IMPORTANT THINGS TO KNOW ABOUT THE JEWISH RELIGION, ITS PEOPLE, AND ITS HISTORY, 62-63 (William Morrow & Co. 1991). On the possibility of viewing oneself from something other than the first person singular standpoint, see Gerald J. Postema, *Morality in the First Person Plural*, 14 LAW AND PHILOS. 35 (1995).

<sup>5</sup> JOHN MARTIN FISCHER & MARK RAVIZZA, RESPONSIBILITY AND CONTROL: A THEORY OF MORAL RESPONSIBILITY 208-10 (Cambridge 1998). See SHARON LAMB, THE TROUBLE WITH BLAMING: VICTIMS, PERPETRATORS, AND RESPONSIBILITY 133-39 (Harvard 1996). Bernard Williams describes this process as one that culminates in the implicit understanding that an agent knows what it means to act “*on reasons*,” a stance that requires a sense of reciprocity which adjudicated wrongdoers, in light of their misconduct, are often unable to manage. BERNARD WILLIAMS, ETHICS AND THE LIMITS OF PHILOSOPHY 65 (Boston: Harvard U. Press 1985). When it comes to blaming children, the point of blaming is the hope of creating self-correcting proleptic mechanisms so that ethical behavior becomes a part of their everyday motivational set. LAMB, *supra* at 133 (describing the process as one of “internalization”). The practice appears to be an emotionally reactive (backwardly directed) expression at conduct (including speech acts) that produce or could produce unwelcome outcomes. See Strawson, *supra* note # –.

<sup>6</sup> Throughout this work, I refer to “capacity-in-fact.” What I mean by the phrase is defined, in part, by its goal: to establish a richer view of human potential than the extant legal view. The view advanced here conduces to treatment and compassion, and measures cognitive and volitional capacities that acknowledge the need for a more exacting assessment of the separation of humankind from lower primates. The new definitions far exceeds the current test of minimal instrumental capacity, which measures only the ability to effect a simple syllogism. See e.g., Michael Moore, *The Moral and Metaphysical Sources of Criminal Law* in J. Roland Pennock and John W. Chapman, eds. NOMOS XXVII: CRIMINAL JUSTICE 20 (New York: N.Y.U. Press, 1985). The existing standard measures only the capacity to move from *desire* to *belief* to *action*, which satisfies the requirements of a practical syllogism. In this respect, we are barely distinguishable (if at all) from lower primates: “Primary consciousness – The ability to generate a mental scene in which a large amount of diverse information is integrated for the purpose of directing present or immediate behavior – occurs in animals with brain structures similar to ours.” Gerald M. Edelman and Giulio Tononi, A UNIVERSE OF CONSCIOUSNESS: HOW MATTER BECOMES IMAGINATION 103 (Basic Books 2000). See also Joseph LeDoux, THE EMOTIONAL BRAIN: THE MYSTERIOUS UNDERPINNINGS OF EMOTIONAL LIFE 126 (Simon & Schuster 1996). When we begin to refine our understanding of human intentionality, we necessarily move toward an appreciation of the benefits of compassion. On the need for a view of rationality that includes empathic capacity, see Stephen J. Morse, *Rationality and Responsibility*, 74 S. CAL. L. REV. 251, 264-66 (2000) (arguing for a “generic partial excusing condition” for those who lack a well developed conscience and capacity for empathy).

development we view incompletely, with snapshots rather than lifelong motion pictures – are qualitatively different practices than those in which we instruct our own children, although they are still fully human practices. With respect to these *others*, we tend to blame far too readily and too harshly. For example, when we blame children to whom we offer unconditional love, our misjudgments are almost always forgivable; these are *our* children and, for that reason if no others, we are almost always forgiven too. The effects of errors in the judgments we make about others, however, should compel us to exercise substantially more precision in the blaming process because these decisions often culminate in harsh, lasting punishment: lost property and liberty and the pain attendant thereto. This need for greater care and precision also obtains because our collisions with the lives of the *others* are almost always sudden, almost always unbidden, and so always epistemologically incomplete; and from the blamer’s perspective, at least, they have lasting effects.<sup>7</sup> These attributions of blame lack the loving judgments we offer our children as we teach them our norms. These *others* are usually total strangers, a substantial number of whom, owing to bad moral luck,<sup>8</sup> reside in a province that lies outside the reach of the learned, ethically-anticipatory mechanisms we routinely teach our kids:<sup>9</sup> lessons we (should) envelope more in love than anger. Given these prominent deficits, if the practice of blaming the many others in our culture operated as it should, with a genuine compassion that recognizes that these others have often lived tragic lives and with the hoped-for aim of moral and ethical development, we would blame wrongdoers if, but only if, (a) we had more nearly complete information – a time lapse portrait instead of a quickly taken Polaroid – covering the antecedent conditions and events that preceded the wrongful act,  $\Phi$ ; (b) we were certain that the wrongdoers had the capacity-in-fact and opportunities to learn the need to avoid inflicting harm on others; and (c) we were confident that we had genuinely tried to set aside our own prejudices in the process of ascribing blame.<sup>10</sup> We would, in short, demand for all

---

<sup>7</sup> For a real life, often wrenching description of how this process works with children adjudicated guilty of violent crimes, see Mark Salzman, *TRUE NOTEBOOKS* (Knopf 2003) (documenting the author’s two years as a creative writing teacher in a youth detention center in Los Angeles).

<sup>8</sup> See, e.g., Thomas Nagel, *Moral Luck*, in *MORTAL QUESTIONS* 24, 34 (Cambridge 1979) (discussing four species of moral luck, including *consequential luck*, *constitutive luck*, *antecedent circumstances*, and luck in the *testing circumstances* one has or has not confronted); Bernard Williams, *Moral Luck*, in *MORAL LUCK: PHILOSOPHICAL PAPERS, 1973-1980*, 29 (Cambridge 1981) (rejecting the Kantian hypothesis that good or bad will is “unconditioned” [or] . . . free from external contingency”); Bernard Williams, *Moral Luck: A Postscript*, in *MAKING SENSE OF HUMANITY AND OTHER PHILOSOPHICAL PAPERS, 1982-1993*, at 241, 241 (Cambridge 1995). *Constitutive luck*, which speaks directly to one’s personality and character, is discussed *infra* text accompanying notes –.

<sup>9</sup> See Williams, *Internal Reasons*, *supra* note # –, at 43 (describing blaming as an advising or “proleptic” mechanism).

<sup>10</sup> On Rawls and the veil of ignorance as a heuristic for achieving distributive justice, see John Rawls, *A THEORY OF JUSTICE* 17-22, 136-42 (Harvard U. Press 1971); *POLITICAL LIBERALISM* (Columbia U. Press 1993); and *POLITICAL LIBERALISM* 22-8 (Columbia U. Press 1993). See generally, Jean Hampton, *POLITICAL PHILOSOPHY* 137-38 (Boulder: Westview, 1997) (noting that the purpose of the “veil of ignorance” is to place those making decisions within a position ensuring that “each person in this position [is] ignorant of her place in society, her particular religious or metaphysical views, her moral beliefs, her social theories and so forth”). I think

children that which we insist upon unwaveringly for our own. That we're not close to meeting these requirements as a matter of routine criminal justice administration (and seem sometimes to lack the will even to question the settled ways<sup>11</sup>) requires us to rethink the role of blaming as a practice in our criminal law.

This rethinking is especially necessary because the social practice of blaming that we pursue today expresses a view of moral and social psychology premised in too large a part on the wrongdoer's wicked *character*. In this work, I examine this practice, beginning with a prominent view from the moral philosophy of blaming, which stresses the bad character of the wrongdoer. Next, I look at the semantics of the term *character* that support blaming generally; when we unpack the term, we discover an understandably abbreviated but, *just to that extent*, woefully incomplete shorthand that conceals deeper problems. And, finally, I discuss some of those deeper problems, which often show up as reflections of the social and cultural biases that we bring to the process of attributing blame. I present a synthesis on the practice of blaming from rich data in the social sciences signaling that our blaming practices tend project our own dispositions for over-emphasizing the actor's wicked disposition to wrongdoing in a process that, on balance, doesn't serve our long term goal of providing a safer society for ourselves and our children: by creating conditions that permit us to over-focus substantially on the *bad* person rather than the situation, by providing impetus to an unhelpful penchant for scapegoating, and, foremost, by undermining our own moral and ultimately political responsibility for the world. Finally, I address blaming as part of the paradox of evil and present suggestions for modifying the extant practices.

## II. A Theory of Blame

One way to pursue issues of moral psychology is to examine a “psychological phenomenon, and explore its links to normative concepts and assumptions.”<sup>12</sup> As Jay

---

this requirement is every bit as important for our criminal justice system as it is for our conception of distributive justice. See, e.g., James P. Sterba, *Retributive Justice*, 5 POL. THEORY 349, 353 (1977) (arguing that the “ignorance condition [might] guarantee fair decisions among persons with conflicting interests. . . . [because] fairness requires that we discount our own special interests in arbitrating conflict-of-interest situations, and imagining oneself to be ignorant of those special interests has the same effect”). I have embraced this approach in other works, including *An Essay on Liberalism and Public Theology*, 14 J. LAW & RELIG. 229 (Winter 1999/2000).

<sup>11</sup> Compare Sanford H. Kadish, *Forward: The Criminal Law and the Luck of the Draw*, 84 J. CRIM. L & CRIMINOLOGY 679, 690 (1994) (“The settled moral understanding is that what you deserve is a function of what you choose.”) with Sanford H. Kadish, *Moral Excess in the Criminal Law*, 32 McGeorge L. Rev. 63 (2000) (disproving of calls to account formally for faultless misconduct attributable to systemic – but currently non-pathological – deficits). Accord, e.g., Michael S. Moore, *Causation and the Excuses*, 73 CAL. L. REV. 1091, 1092 (1985) (arguing that “moral responsibility for an action should be ascribed to an actor even when that action was caused by factors over which he has no control”).

<sup>12</sup> R. Jay Wallace, *Moral Psychology*, in Frank Jackson and Michael Smith, eds., THE OXFORD HANDBOOK OF CONTEMPORARY PHILOSOPHY (Oxford University Press) (page cites unavailable), at [www.ist-socrates.berkeley.edu/~jwallace/papers/rtf/Moralpsych.rtf](http://www.ist-socrates.berkeley.edu/~jwallace/papers/rtf/Moralpsych.rtf).

Wallace notes, “If we wish to understand morality as a normative phenomenon, we must understand the psychological states and capacities that make it possible for it to regulate people’s lives in the way it aspires to do.”<sup>13</sup> The question here is how does the practice of blaming connect complex psychological states with developable capacities, on the one hand, and, on the other, how do both connect with the norms to which we aspire as a compassionate society. This section begins with the views of a preeminent moral philosopher, Judith Jarvis Thompson, and then turns to the work of several contemporary jurisprudence scholars who advance character-based theories of punishment and excuse. On the (more or less) conventional view of moral psychology explored here, an actor is “bad” if he acts in ways that (a) make his conduct more blameworthy than that of a lucky someone else whose intentions are indistinguishable but produce no harm, and (b) cause us to believe that his unwelcome conduct reveals an essential deficit in his character. It then moves on to consider the merits of this general approach in light of well-replicated social science data on blaming as a social practice.

Thompson presents her approach to our understanding of blame in an essay on the effects of moral luck. She hopes to demonstrate that bad luck concerning one’s antecedent and constitutive circumstances (roughly genetic and early socio-economic conditions, respectively<sup>14</sup>) does not undermine the possibility of fully warranted individual responsibility.<sup>15</sup> Her analysis is helpful but incomplete, and her rhetoric brings to view some of the unexamined presuppositions that guide her (and us) to the semantics and attributions of blameworthiness. Thompson offers two intention-based definitions of blame in an effort to tame the determinist beast. The first blame formulation, **BF1**, states that *a person is to blame for an unwelcome event or state of affairs where that person caused the event or state of affairs by some wrongful act or omission for which he had no adequate excuse*. The second, alternative formulation, **BF2**, holds that *a person is to blame for doing something to a greater or lesser extent where his doing that something gives us a stronger or weaker reason to think he’s a “bad” person*.<sup>16</sup> **BF2**, unlike **BF1**,

---

<sup>13</sup> *Id.*

<sup>14</sup> See note – and accompanying text *infra*.

<sup>15</sup> Judith Jarvis Thompson, *Morality and Bad Luck*, in *MORAL LUCK* 195 (ed. by Daniel Statman; SUNY 1993). In his work on moral luck, Nagel notes unexceptionally that we commonly assert that “[a] person can be morally responsible only for what he does.” At the same time, however, we know with certainty that “what he does results from a great deal that he does not do.” In the end, then, we face a paradox: The actor “is not responsible for what he is and is not responsible for.” On this view, the actor seems to be morally responsible for nothing, because for much that he does he bears no responsibility – a view which seems to be normatively intolerable despite descriptive accuracy. Nagel, *supra* note # –, at 34.

<sup>16</sup> *Id.* at 200-02. The undefined term *bad* is Professor Thompson’s choice. She offers a variation of this first formula that focuses instead on *being the kind of person who* causes an unwelcome event instead of focusing on causation; that is, a person is to blame for *being* in a certain way that produces unwelcome outcomes rather than *causing* the same. Although this formulation puts character more clearly into focus, I do not think the differences matter for present purposes; in the absence of describing “the kind of person who” in terms that explain *why* he is “the kind of person who,” in her telling “causing the kind of event which” describes the same personality and

stresses the effect on the viewer and raises questions about the compatibilist construct within which Thomson operates.<sup>17</sup>

Thompson's goal is to account for blame in a way that acknowledges causal determinism<sup>18</sup> without succumbing to it.<sup>19</sup> In that spirit, she offers two additional theories to account for (or reject) blaming as a social practice in a world that's causal all the way down. The first, the strong determinist thesis, **SDT**, holds that blaming is always improper because personal responsibility is, by definition, lacking; whatever happens must happen. Thus no misconduct can bring discredit to us.<sup>20</sup> She also offers a weak determinist thesis, **WDT**, according to which events over which we have no control (luck) matter, but that nothing we do in consequence thereof sheds any more discredit on us than is shed on us by our mental acts.<sup>21</sup> **WDT** rests on the reasonably hypothesized conclusion that anytime we

---

character ensemble and set of antecedent events and conditions (or states and capacities). She makes no such distinctions, however.

<sup>17</sup> For a discussion of the distinction between "hard" determinism, according to which freedom is an illusion because *all* of our conduct is *always* determined, and "soft" determinism, according to which one can acknowledge that events have antecedent causes, but maintain nonetheless that in an important sense agents control their own mental states, even though those states are themselves causally determined, see P. S. Greenspan, *Behavior Control and Freedom of Action*, in *MORAL RESPONSIBILITY* 191, 191 n. 1 (ed. by John Martin Fischer; Cornell 1986).

<sup>18</sup> "Causal determinism is the thesis that, for any given time, a complete statement of facts about that time, together with a complete statement of the laws of nature, entails every truth as to what happens after that time." Fischer and Ravizza, *supra* note # –, at 14 (arguing, generally, that causal determinism, whether true or not, does not subvert our common sense understanding of moral responsibility). Although one could distinguish between determinism and mechanism, Hilary Bok, *FREEDOM AND RESPONSIBILITY* 4 (1998), whereby determinism describes one form of mechanism, I use the two terms synonymously here.

<sup>19</sup> This, of course, was a task at which Kant ultimately failed. See H. J. Paton, trans., *KANT'S GROUNDWORK OF THE METAPHYSICS OF MORAL* 117 (New York 1873). Kant embraced the antinomy of practical reason which, famously, he reconciled through the fiction of an inaccessible "thing . . . in itself." T. K. Abbott, trans., Immanuel Kant, *CRITIQUE OF PRACTICAL REASON* 138-45 (Prometheus Books 1996). John Rawls commented that the "weakness" of Kant's abstraction of the foundations of a "moral law" was in its recognizing "only two possibilities: either the moral law is founded on an object given to it," like Hume's conception of the "passions," or "as pure practical reason determines . . . its own object of itself." John Rawls, *LECTURES ON THE HISTORY OF MORAL PHILOSOPHY* 235 (ed. by Barbara Herman; Cambridge, MA: Harvard, 2000) (hereafter "Rawls, LECTURES").

<sup>20</sup> The basic analytic conclusion is that if determinism is correct, then blame is always undeserved. See generally A. J. Ayer, *Freedom and Necessity*, in Watson, *supra* note # –, at 15.

<sup>21</sup> With Adam Smith, who famously discussed the "irregularity of sentiments" in *THE THEORY OF MORAL SENTIMENTS*, *supra* note # –, at 135, Thompson subscribes to the prevailing view that we judge blameworthiness based only on intentions and not consequences. See, e.g., Joel Feinberg, *Equal Punishment for Failed Attempts: Some Bad But Instructive Arguments Against It*, 37 *ARIZ. L. REV.* (1995); Richard Parker, *Blame, Punishment, and the Role of Result*, 21 *AM. PHIL. Q.* 269, 271 (1984) ("The individual is blameworthy and punishable, on my view, only for conduct itself, where conduct is construed as a combination of overt action, state of mind (including intention, knowledge, etc.) and circumstances."). See Sanford H. Kadish, *Forward: The Criminal Law and the Luck*

can imagine a person, P, intentionally doing  $\Phi$ , which causes a bad outcome, O, we can also imagine another person, P<sup>2</sup>, undertaking the same  $\Phi$  but, for a variety of reasons, P<sup>2</sup> *fortunately* fails to produce a negative result,  $\emptyset$ .<sup>22</sup> Accordingly, she concludes that we should blame P for doing  $\Phi$  (where the unfortunate  $\Phi$ ing was running over a child when P backed out of the driveway without looking) but no more than we blame P<sup>2</sup> (who also  $\Phi$ ed when he backed out of his drive without looking but had the good fortune that no unseen child was crossing his vehicle's path); both acted "wrongfully" although P<sup>2</sup> happily avoided the bad outcome.<sup>23</sup>

Thompson rejects *SDT*, which she rightly explains as holding that "whatever . . . [P or P<sup>2</sup>] does, the event that consists in his or her doing it has to occur" – a hard determinist position that undermines moral responsibility entirely.<sup>24</sup> Under this view of the world, Thompson argues, the person who deliberately kills babies for amusement suffers no discredit and no blame. She disavows this view because, in her words, "*of course*" the baby killer, BK, is a "*bad person*."<sup>25</sup> *SDT* can't be true "if it's taken to deny the second notion of 'blame,'" according to which we have good "*reason to think [BK] is a bad person*."<sup>26</sup> Thompson doesn't deny that luck plays a formative role in who the actor is at *t*, the

---

*of the Draw*, 84 J. CRIM. L & CRIMINOLOGY 679 (1994) (arguing that the harm doctrine, whereby punishment varies based upon the chance that harm does or does not occur, is rationally indefensible). I have taken up professor Kadish's challenge in *A Jurisprudence for Punishing Attempts Asymmetrically*, 6 BUFF. CRIM. L. REV. 951 (2003).

<sup>22</sup> In this context, "intentional" refers broadly to our mental interactions with and about the world, not to the narrow description of statutory criminal mens rea. See, e.g., Model Penal Code §2.02 (2) (American Law Institute 1985) (defining levels of criminal intent including negligence, recklessness, knowledge and purpose) (hereafter, cited as "M.P.C.")

<sup>23</sup> Although she purports to be offering Thomas Nagel's explanation of moral luck, she takes it on as her own in an effort to challenge the notion that ordinary events undermine moral responsibility. Hence she writes frames the challenge this way: "Take any person who is engaging in *X* caused outcome *O*. We can always imagine a second person engaged in *X*, but whose engaging in *X* did not cause an outcome of kind *O*." *Id.* at 199. The "wrongful" conclusion is, of course, not incontestable from the perspective of tort law, as the debate between Holmes and Andrews makes clear. Compare *Palsgraf v. Long Island R.R.*, 162 N.E. 99 (Ct. App. N. Y. 1928) (Holmes, J.) with *id.* at 104 (Andrews, J., dissenting).

The theory of action she follows, which accounts for our intentions and potential fault, is fairly unexceptional, although admittedly thin: Whatever a person does is either a purely mental act, or is a causing of some outcome by some mental acts or acts—and that something can be either an act or an answerable omission. On this account of action, if one forms the intention of washing clothes in the washing machine and does so, putting the clothes in the machine was causing the clothes to get washed, which was a product of one's mental acts, among which was the intention to wash clothes. *Cf. id.* at 199.

<sup>24</sup> Thompson, *supra* note # – , at 208.

<sup>25</sup> Thompson, *supra* note # – , at 208 (emphasis added).

<sup>26</sup> See *id.*



theoretical moment when bad act and criminal intent converge.<sup>27</sup> It is this acknowledgment that leads her to **WDT**, a view of determinism that permits us to measure blame by those purely mental acts that produce the misconduct or a chargeable omission. **BF2**, because it is only comparative – Do we have a stronger or weaker reason to believe P is a bad person compared to P<sup>2</sup>? – is nonetheless inadequate. Instead of addressing **BF2**'s inadequacy directly, however, she lashes the two formulations of blame to a thin concept of an adequate excuse.

She accomplishes this tie-in by constructing a theoretical compatibilist link between **BF1** and **BF2**, which she labels a “connecting thesis” (**CT**). **CT** states that *a person is to blame for some conduct Φ under BF1 (causing an unwelcome event) if and only if he is in some way to blame for Φing under definition BF2 (causing us to think he is a “bad person”) for which there is no adequate excuse.*<sup>28</sup> Thus the wanton baby killer, BK, is blameworthy under **CT** because he caused an unwelcome event that causes us to think he is a bad person, and he has no excuse for the conduct “since his [killing babies] *does* give us reason . . . to think him a bad person.”<sup>29</sup>

No one doubts that BK is a “bad” person, as a matter of ordinary usage the world over. Austin noted years ago that “ordinary language is not the last word: in principle it can *everywhere* be supplemented and improved upon and superseded;” it is importantly, however, “the *first* word.”<sup>30</sup> Because it only reflects our pre-evaluative intuitive beginning; it is defeasible. **CT** is not the last word. Thompson’s pervasive use of the descriptor, *bad*, raises at least three questions: (a) What is an acceptable definition of *adequate excuse* so that what at first appears to be *bad* conduct is excused? (b) Do her hypothetically inadvertent and morally indistinguishable identical wrongdoers, whose conduct provides the model for her theory of blame, really apply to brutal BK, despite their different levels of criminal intent? Most importantly for present purposes, (c) What is entailed in concluding that a person is *bad*? This last question introduces the character theory of jurisprudence and conditions that define criminality.

---

<sup>27</sup> See, e.g., *Morrisette v. United States*, 342 U.S. 246, 251 (1952) (“Crime, as a compound concept, generally constituted only from concurrence of an evil-meaning mind with an evil-doing hand.”). The Model Penal Code reflects the prevailing wisdom by requiring convergence of a voluntary act with the statutorily prescribed level of culpability. See M. P. C. §§ 2.01(1), 2.02(1). Mark Kelman, *Interpretive Construction in the Substantive Criminal Law*, 33 STAN. L. REV. 591 (1981), has demonstrated, however, that even this bedrock requirement is subject to manipulation: the courts often expand or contract the relevant time frame for this convergence to occur as a matter of public policy. See, e.g., *People v. Watson*, 637 P.2d 279 (Cal. 1981) (affirming a conviction for second degree murder where the defendant knew several hours earlier that he would be driving his car when he left the bar); *People v. Decina*, 138 N.E. 2d 799 (N.Y. 1956) (overruling a demurrer to the charge of culpable negligent homicide where the defendant, “*knowing*” that he was subject to seizure disorders, chose to drive his car, lost consciousness, and caused several deaths).

<sup>28</sup> Thompson, *supra* note # –, at 211.

<sup>29</sup> *Id.* at 211-12.

<sup>30</sup> J. L. Austin, *A Plea for Excuses*, 57 PROCEEDINGS ARISTOTELIAN SOCIETY 1, 11 (1956).

One response to (a) is that she leaves substantially undefined the components of an adequate excuse. She defines excuse broadly (and admittedly tentatively) as a condition or event that annuls **BF2**, conduct that causes us to label the actor “bad.” Even this hasty definition is mere tautology; it tells us nothing about the wrongdoer’s capacity-in-fact to control his mental state.<sup>31</sup> Generally, an excuse annuls the wrongdoing because the actor is, for some reason, not blameworthy: His conduct reflects some incapacity such that  $\Phi$ ing is viewed as involuntary or unknowing or unintentional. Excuses typically speak to involuntariness, not in the sense of the absence of a bodily act, but in the sense of absence of genuine choice, which “does not accurately reveal the actor’s character,” and thus makes punishment unjust.<sup>32</sup> In the case of blaming BK, often the first issue in question should be one of capacity-in-fact, a question philosophers tend not to pursue.

Thompson’s slender definition of an “adequate excuse” is not the only problem. One additional difficulty concerns her effort to construct hypothetical examples that reflect varying levels of intentionality that, in turn, produce equal culpable states between actions and omissions among her paired, hypothetical wrongdoers; another addresses her unanalyzed choice of metaphors. The first responds to the hypothetical situations she constructs as a platform for her illustrations of comparative blameworthiness; they are too imprecise to support her all-purpose notion of *badness*. Her hypothesized examples of wrongdoing are intended to illustrate levels of individual blameworthiness that are, from the standpoint of moral accounting, comparatively indistinguishable across the broad expanse of culpable intent, and this project would fulfill the demands of *CT*. Her examples fall short of supporting her goal; their basic impulse is far too formal.

She begins convincingly by demonstrating that the hapless motorist, P, whose  $\Phi$ ing consists of failing to look into his rearview mirror before backing out of the drive, is neither more nor less guilty of wrongdoing than lucky P<sup>2</sup>, whose identical  $\Phi$ ing luckily inflicts no harm. Her approach at the inadvertence end of the intent continuum is well taken; we all suffer from self-condemnation and with corresponding feelings of self-blame (but usually less, in varying degrees, than it might have been if the harm had unfortunately come to pass).<sup>33</sup> Causing pain to others does and should bring pain to us even when the act

---

<sup>31</sup> In fairness, I do not know what Professor Thompson’s thoughts are about the presumptions of sufficient capacity for control. She does not qualify her observations with this question in mind. For a helpful approach in a book length treatment, see Bok, *supra* note # –, at 7, 8-9 & n. 10, who offers a comprehensive theory for acknowledging free will but states at the outset that she is leaving some questions unaddressed, including the moral responsibility of the addict who continues to abuse her drug of choice, of those who act under compulsions or “powerful unconscious motives,” and issues of punishment.

<sup>32</sup> George Fletcher, *The Individualization of Excusing Conditions*, 47 S. CAL. L. REV. 1269, 1271 (1974). For a theory of excuses tied to the issue of choice, see generally Sanford H. Kadish, *Excusing Crime*, 75 CAL. L. REV. 257 (1987).

<sup>33</sup> This is a point Smith fully elucidated in *THEORY OF MORAL SENTIMENTS*, *supra* note # –, at 134 (concluding that the “intention or affection of the heart” is all that matters to assess blameworthiness, but that in “particular cases the actual consequences which happen to proceed from any action, have a great effect upon our

producing the harm is unintentional and non-malevolent, albeit insensitive.<sup>34</sup> Her argument becomes more difficult, however, when she transports a description of blameworthiness created in the context of empty-headedness to the context of baby killing; the relative blameworthiness of these two instantiations of intention is unstable. No longer are we assessing the blameworthiness of an empty or other-directed but relatively normal mind; with BK we are viewing the impact of a mind filled with crazed deliberations of the most motivationally repugnant even unimaginable form: pain for its own sake or, equally deplorable, without any concern at all about the harm imposed relative to some perverse immediate benefits gained.<sup>35</sup> When we describe a person as *bad* in the context of inadvertence, we are mostly describing the unwelcome effect of oversight. The examples she uses initially to establish *WT* arise in situations where P and P<sup>2</sup> have the same legal duties to avoid  $\Phi$ ing: that is, to keep a careful lookout. Thus the description of their misconduct as *bad* speaks mostly to the effect of their breach of duty and not (necessarily) to their general characters or dispositions;<sup>36</sup> for there but for the grace of God go any one of us. In stark contrast, when we talk about *BK*, we are addressing a qualitatively different species of wrongfulness.

---

sentiments concerning . . . merit or demerit, and almost always enhance or diminish our sense of both”). I have discussed this at length in terms of the aesthetics of retribution in *Aesthetics of Retribution*, *supra* note # at 243-47 (surveying some of the implications of the positions of Smith and Hume on the normativity of a naturalist approach to moral philosophy).

<sup>34</sup> See Andre, *supra* note # –.

<sup>35</sup> Cf. R. A. Duff, *Choice, Character, and Criminal Liability*, 12 L. AND PHIL. 345, 348 (1993) (stating that liability for criminal negligence does not require choice unless negligence is defined in the pluperfect – she would have chosen to avoid do  $\Phi$  if she had made a choice). And that conclusion would beget a “recklessness” charge from many quarters.

Nor does a classic utility approach, which finds rationality in the ability to determine one’s own immediate hedonic needs. On the classic side, see Bentham, *Principles of Penal Law*, *supra* note # –, Ch. XV, §1, at 83, who stated in summary form:

- (a) All laws should promote happiness, i.e., net social gain;
- (b) Punishment itself does not promote happiness;
- (c) Therefore, laws requiring punishment should be enforced only when doing so otherwise promotes happiness, i.e. produces a net social gain.

One need not assume that utility, defined as aggregate happiness, is purely hedonic. See, e.g., John Rawls, *Two Concepts of Rules*, 44 PHIL. REV. 3, 9 (1955) (amending the utility principle to make clear that it is *not* purely hedonic, and thereby providing utilitarian theory with a rationale that would prevent punishing the innocent).

Finally, there is no neurobiological support for a distinct human-ness that guarantees civility. See note – *supra* (discussing the phenomenon of “primary consciousness”).

<sup>36</sup> In fact, Thompson goes to some lengths to describe P, P<sup>2</sup>, and P<sup>3</sup> (also inadvertent) as generally careful drivers who were momentarily distracted. Thompson, *supra* note –, at 196-98.

Thompson does offer a mid-level, apparently transitional example of bad intent (between a motorist's inadvertence and perverse baby killing) when she includes within her moral model the mental acts of two hypothetical judges, J and J<sup>2</sup>, both equally prepared (and willing) to accept a bribe and undermine justice. This reflects a level of intent akin to "knowledge" at least.<sup>37</sup> Of the two, only J is actually approached with and offered a bribe; J accepts. She argues, again, that J<sup>2</sup>'s mental acts makes him equally as culpable as J because J<sup>2</sup> also was willing to accept the bribe. This example of advertent, intentional wrongdoing moves her thesis closer to a general approach that supports **CT** than does the example of the distracted neglectful drivers, but I think the case of the corruptible judges also presents a different characterological matter than does the case of the gleeful baby-killers. The case of J and J<sup>2</sup> illustrates uncommon but motivationally intelligible greed, a common form of culpability and a tendency most people have had to tame at one time or another. The comparison with the enraptured Jeffrey Dahmer-like,<sup>38</sup> motivationally unintelligible butchery fails, however, because it raises all sorts of psychopathological and neurobiological issues that the conduct of J and J<sup>2</sup> does not implicate. We more clearly understand J than we do BK.

Baby killing as sport is *for any reason* "bad" beyond the imaginations of most of us. It is the unfathomable taking of a human life – someone's cherished, innocent child – for no reason, not even an explicable Andrea Yates-like psychotic reason.<sup>39</sup> The baby killer

---

<sup>37</sup> See M. P. C. §2.02 (2) (a & b) (distinguishing between conduct that is undertaken with either the purpose – "conscious object" – or the knowledge – "practical certainty" that prohibited conduct or results will occur).

<sup>38</sup> See, e.g., *Insanity Defense By Dahmer Fails*, N. Y. TIMES, Feb. 16 1992, sec. 1, part 1; col.6; *Report Documents Dahmer's Cannibalism*, U.P.I., Aug., 4, 1991; online at [www.geocities.com/donnysmith.geo/dahmer.html](http://www.geocities.com/donnysmith.geo/dahmer.html); [www.arminm.com/jeffrey\\_dahmer.htm](http://www.arminm.com/jeffrey_dahmer.htm).

<sup>39</sup> Andrea Yates is the Houston mother who drowned her five young children with the hope that doing so would provide for their salvation. See *Transcript of Andrea Yates' Confession*, Feb. 21, 2002 at [www.chron.com/cs/CDA/story.hts/special/drownings](http://www.chron.com/cs/CDA/story.hts/special/drownings). On her background and those of her husband and "spiritual advisor," see *A Mother's Madness - Andrea Yates*, at [www.courtvtv.com/onair/shows/mugshots/episodes/yates.html](http://www.courtvtv.com/onair/shows/mugshots/episodes/yates.html).

In brief, Yates was described by friends as a "wonderful mother;" they didn't know how hard she struggled with debilitating depression, a depression that produced psychotic delusions. Andrea Yates suffered from "postpartum depression," a condition which, in less severe forms, effects many women after childbirth; in severe forms, it can induce psychotic states. She killed all five of her young children, one at a time, by holding their heads under water in the same family bathtub where she had bathed them during the days and weeks and months and years of ignored warnings that preceded her fatal crisis. A devout fundamentalist Christian, she followed her husband's and their spiritual adviser's mantra – "God will take care of us!" – and had yet another child despite increasingly difficult bouts of depression and decomposition after the births of her first four children. Following the birth of their fifth child, she came to the bizarre conclusion that she (now clearly losing her struggle with sanity) was possessed by Satan and was causing her children irreparable and eternal damage. Guided by the singular logic of her unique psychosis, she decided that the only chance they had of getting to heaven required her to kill them one by one, a task she completed despite considerable struggle. The examining psychiatrist agreed with the state that at least in some moments, Yates knew that it was legally wrong to murder her children, but she believed she had no choice: "Because of her dilemma, what she perceived as right was to take her children's [sic.] life on earth to prevent them

causes great (if brief) pain to the deceased child, and massive theretofore unimaginable, life-changing pain to the baby's parents, families, and friends. This unvarnished description of pain produces an emotional aesthetic that is often uncabinable and immeasurable.<sup>40</sup> Here the context for Thompson's definitions of blame blurs important distinctions. First, unlike P and P<sup>2</sup> (the negligent drivers) or even the corruptible judges, J and J<sup>2</sup>, it is difficult even to *imagine* a BK<sup>2</sup> for every BK; that is, we can only barely imagine a BK, much less BK<sup>2</sup>, whose misconduct consists in only *attempting* to kill babies for fun; and unlike the distracted driver or the greedy judge, we cannot imagine at all what motivates these hypothetical baby killers.<sup>41</sup> Second, excuses speak only indirectly to the *bad* label we place on the actor's conduct – that is, to whether we have a stronger or weaker reason to think he's a bad person. Rather, they speak to the actor's mental state, traditionally whether he was ignorant or coerced or mistaken or the like. We excuse him, in part, *because* he was not acting like himself or he was acting involuntarily or he lacked significant cognitive or suffered a significant deficit in volitional control. Short of something like an unforeseen heart attack or other genuinely involuntary act, there simply is no excuse for the inadvertent drivers, P or P<sup>2</sup>.

The same, paradoxically, cannot be said of BK. No one doubts the *badness* of BK's character or of Dahmer's or of the barbarity of their crimes, at least as a matter of common expression. No feeling, caring, even marginally empathic human can harbor such a desire, much less the demand that it be gratified with such gruesomeness.<sup>42</sup> If that conclusion is accurate, and it must be, then we should pursue a deeper question, which Thompson omits: What aberrant synergies of poor environment and faulty neurocircuitry make it possible for any actor to achieve a state of mind that produces insatiable desire and joy from the purposeful act of killing precious innocent babies? So framed, to call the actor *bad* or *evil* is to say, metaphorically, that he is theologically irredeemable, a statement we cannot make about P or P<sup>2</sup>. Thus we should ask the common sense question: How did this person BK get so sick? And if we agree that BK (or Dahmer or Yates and others) are sick, what ingredient in the nature of blaming is served by blaming him? Blaming those who lack the

---

from eternal damnation." According to psychiatric testimony, Yates believed at times that she would sacrifice her life to save her children from damnation. "She was afraid Satan would hear it and make it happen." Yates was deemed "sane;" competency was not an issue.

<sup>40</sup> And so the retributive emotion that's triggered by the idea of frivolous baby killing is undeniable in its existence (even if the conduct Thompson describes is, thankfully, rare.) I take this topic up directly in *Some Thoughts on the Aesthetics of Retribution*, 17 CAN. J. L. & JURIS. 233 (2004). It is true that negligent killings also may evoke retributive emotions but not (ordinarily) vengeful ones; we tend to address negligent deaths through money damages rather than incarceration or worse.

<sup>41</sup> Moreover, unlike the careless but fortunate driver, with whom we can all sympathize *and* advance some reproach, there is no deliberate baby killer who has the good fortune of not killing – or at least of causing grave harm to – an infant he hopes to kill, unless BK<sup>2</sup> is caught prior to inflicting harm.

<sup>42</sup> The account of reactive attitudes I describe below begins with Strawson, *supra*, note # –. For chilling portraits of psychopathic personalities, see R. D. Hare, *WITHOUT CONSCIENCE: THE DISTURBING WORLD OF THE PSYCHOPATHS AMONG US* Ch. 1 (New York: Simon and Schuster, 1993).

capacity-in-fact for moral understanding is pointless, not in all instrumental senses because, as noted above, we do blame our children when we ascribe accountable fault to them for producing an unwelcome event of state of affairs. Rather, to blame BK is an empty gesture inasmuch as the point of blaming is to communicate resentment; there is no communication possible with certain people,<sup>43</sup> at least without an effort at offering substantial long term assistance, or killing. Why then do we so quickly determine *badness* without considering the conditions and events – the whole biographies, insofar as that information is available – of actors who behave badly? What conditions and events produced these wretched characters?<sup>44</sup>

The deficiencies they experience combine unfortunate events with unfortunate conditions.<sup>45</sup> Reflect momentarily on the background of Andrea Yates, who suffered from a psychotic state that was induced by repeated, tragically under-examined bouts of postpartum depression;<sup>46</sup> or on Robert Alton Harris's, whose childhood was as devoid of love as it was occupied with and impacted by hate and abuse;<sup>47</sup> or on Jeffrey Dahmer's, whose genetic defects caused him to commit totally unfathomable acts: carving, storing, and cannibalizing his victims. These tragic figures could not have caused their own deranged characters, at least not in any identifiably linear fashion. In fact, if we accept the determinist premise that Thompson hopes to corral, that every event has a cause and that

---

<sup>43</sup> Gary Watson, *Responsibility and the Limits of Evil*, in RESPONSIBILITY, CHARACTER, AND THE EMOTIONS, 265 (ed. by Ferdinand Schoeman; Cambridge 1987) (hereafter, "Watson, *Responsibility*").

<sup>44</sup> There might yet be a certain population against whom we cannot defend ourselves, see [ARYAN NATION ARTICLE]; and as to their fate and its moral metric, I do not know..

<sup>45</sup> See Matt Ridley, NATURE VIA NURTURE: GENES, EXPERIENCE, & WHAT MAKES US HUMAN (HarperCollins 2003) (describing the we are genetically coded to responded to environmental stimuli) [hereafter, "NATURE VIA NURTURE"].

<sup>46</sup> See Sheri L. Beinstock, *Mothers Who Kill Their Children Postpartum Psychosis*, 32 SW. U. L. REV. 451 (2003).

<sup>47</sup> On Harris's life, see Miles Corwin, *Icy Killer's Life Steeped in Violence*, LOS ANGELES TIMES, May 16, 1982 (reprinted in Watson, *Responsibility*, *supra* note # –, at 268-71, 272-74). Corwin presents the following portrait: Harris was the product of cruelly abusive chronic alcoholic parents. His mother not only refused to permit him to make any physical contact with her ever, she blamed him, the youngest of nine children, for all her suffering. His father, never willing to accept that his youngest child Robbie was his progeny, kicked his mother in the stomach while the child was *in utero*. Later that same night the child was born and nearly died. His sister recalled that as a beautiful young child all Robbie had ever wanted was love from someone – anyone. He never received it. Twenty five years after his rude birth, Harris brutally and senselessly murdered two young teens and then calmly ate their lunches. His fellow death row inmates were so disconcerted by Harris's demeanor that they cheered his impending death.

On the issue of environmental effects, see Ridley, NATURE VIA NURTURE, *supra* note # –. See also Antonio R. Damasio, DESCARTES' ERROR: EMOTION, REASON, AND THE HUMAN BRAIN (HarperCollins 1994) (discussing the somatic marker thesis); Joseph LeDoux's, SYNAPTIC SELF: HOW OUR BRAINS BECOME WHO WE ARE (2002).

every cause has a prior cause all the way down, the very notion of self-causation is either absurd or meaningless. Galen Strawson explains why:

[T]o be truly deserving of . . . praise or blame for our actions, . . . we must be truly responsible for how we are mentally [for our intentions.] We must be genuine ‘originators’ of ourselves. . . . But the attempt to describe how we could [possibly] be true originators of ourselves . . . leads to infinite regress . . . for even if one could somehow choose how to be, . . . one would . . . already have to have existed prior to that choice, with a certain set of preferences about how to be. But then the question would arise: where did these preferences come from? Or were they just there, unchosen preferences for which one was not responsible? And so on.<sup>48</sup>

As the next section indicates, self-causation, Aristotle’s hook on character and its impact, ultimately makes no sense. And still, the irregularity of our sentiments being what they are, *we* blame Yates and Harris and Dahmer and BK for their demented characters. Why? Clearly, neither determinism nor indeterminism as heuristic categories provides a dispositive answer. In fact, when we contemplate the breakdown that occurred in the lives of these individuals, we should be concerned with determining the sources of their crazy preferences. There may be no complete answers to this inquiry, but I think there are at least some good questions, including why we blame those who often lack the necessary capacity-in-fact for volitional control.<sup>49</sup>

### III. Character and the Semantics of Blame

What do we think we mean, and what do we have reason to know we mean, when we say that an actor is “bad” or that he has a bad character; that he deserves our condemnation; and, therefore, that he should be blamed and punished for his wrongdoing? One response is that he has done a harmful or potentially harmful unwelcome act that causes us to believe he is a *bad* person. Thompson’s conclusion is analytically objectionable and ontologically incomplete. Many commentators go beyond Thompson’s modest proposal, which stresses a helpful regime of normatively comparative culpability judgments, and insist that we should punish actors only (and always) when their misconduct and subsequent bad choices

---

<sup>48</sup> Strawson, *Consciousness, Free Will*, *supra* note # –, at (1989). *Accord e.g.*, Thomas Nagel, WHAT DOES IT ALL MEAN? 55-7 (Oxford 1987); Hilary Bok, FREEDOM AND RESPONSIBILITY 201 (1998) (“The supposition that we were caused to choose as we did by something other than ourselves implies that we are not free. The supposition that we caused them to occur simply pushes the problem back.”); Stephen J. Morse, *The Moral Metaphysics of Causation and Results*, 88 Calif. L. Rev. 879, 881 (2000) (“No one is a prime mover unmoved.”).

<sup>49</sup> Compare Robert McAfee Brown, *Introduction*, in Elie Wiesel, THE TRIAL OF GOD xvi (New York: Schocken Books, 1995) (“I do not have any answers,” Wiesel acknowledged, “but I have some very good questions.”).

reflect a defect of character (or dispositions or personality).<sup>50</sup> This issue is both chronic, inasmuch as it implies the existence of a non-controversial rule or standard of measurement, and always acute because for many people the conclusion that a wrongdoer is “bad” or “just bad” provides a sufficiently accurate description of character to justify not only blame but imprisonment: locking up defendants who have been adjudicated guilty and throwing away the key. There are at least two replies to the questions that begin this paragraph: one combines a contemporary rule utilitarian outlook on excuses with a thin formal Kantian view of capacity-in-fact; the other proceeds cautiously, responding to widespread social psychological and semantic phenomena entailed in our practice of blaming. I address the two responses in that order, and then take up an issue hinted at and implied throughout my critique: the paradoxical nature of evil and its relation to the regime of excusing misconduct.

A. A brief look at the jurisprudential account of character – In his *ETHICS*, Aristotle insisted that individuals are responsible for their own characters because their actions establish their character.<sup>51</sup> Hume advanced a related idea in the *TREATISE*: “Actions are by their very nature temporary and perishing; and where they proceed not from some cause in the characters and disposition of the person, who perform’s them, they . . . can neither redound to his honor, if good, nor infamy, if evil.”<sup>52</sup> He elaborated the dispositional viewpoint with an example of homicide.

Take any action allow’d to be vicious: Wilful murder, for instance. Examine it in all lights, and see if you can find that matter of fact , . . . which you call *vice*. In which-ever way you take it, you find only certain passions, motives, volitions and thoughts. There is no other matter of fact in the case. The vice entirely escapes you . . . till you turn your reflexion into your own breast, and find a sentiment of disapprobation, which arises in you, toward this action. It lies in yourself, not in the object.<sup>53</sup>

Hume’s view is that we disapprove of BK’s conduct not because it is vicious; rather, we identify it as vicious because we, in turning to our own dispositions (our “breasts”),

---

<sup>50</sup> For useful surveys and critiques of character theories generally, see R. A. Duff, *Do We Want an Aristotelian Criminal Law?*, 6 *BUFF. CRIM. L. REV.* 147 (2002); Duff, *Choice, Character, and Criminal Liability*, *supra* note # –.

<sup>51</sup> Aristotle, *NICOMACHEAN ETHICS* iii.5.1114a3-11 (Trans. by David Ross Oxford, 1987), maintained that by virtue of their slackardly dispositions, men are “responsible for becoming men of that kind.” For a neurologically-informed critique of this view, see Carl Elliott, *THE RULES OF INSANITY: MORAL RESPONSIBILITY AND THE MENTALLY ILL OFFENDER* 29-31 (Suny 1996) (noting that “in many ways we are clearly *not* responsible for our characters, at least as Aristotle implies”). *Id.* at 30.

<sup>52</sup> David Hume, *A TREATISE OF HUMAN NATURE* 411 (L. A. Selby-Bigge and P. H. Nidditch, eds.; Oxford 2nd ed. 1978). On Hume’s particular brand of utilitarian thinking, see Paul Russell, *Hume on Responsibility and Punishment*, 20 *CAN. J. PHILOS.* 539 (1990).

<sup>53</sup> *Id.* at 468-69.



disapprove of it.<sup>54</sup> Morality, according to the Humean view, resides in us, in our characters and dispositions. Such a determination would be relatively unimpeachable if all who are judged bad possessed control-in-fact over their characters.

In his oft cited work, *A Utilitarian Theory of Excuses*, Richard Brandt follows this Humean line, arguing for a rule-utilitarian theory of excuses according to which moral demerit is judged by the following dictum:<sup>55</sup> “An agent is morally blameworthy . . . for an act if, and to the degree that, the moral code the currency of which in that society would maximize utility would condemn . . . him for it.”<sup>56</sup> Condemnation is appropriate, on this view, whenever the actor lacks an internalized motivational system that should produce appropriate feelings of guilt or shame or dishonor and the like in response to conduct that violates the moral code, absent a traditional assessment of excuse. Motivation is key; an agent’s conduct is excusable, therefore, if but only if, “an objectively wrong action (or an action in some way out of order) . . . does not manifest some defect of character.”<sup>57</sup> Brandt’s description of character, which he distinguishes from “just a trait of personality,”<sup>58</sup> is terse; it is some internal quality that goes beyond a trait to include more or less permanent dispositions, so that for Brandt the crucial moral question is whether or not people *would* be “trained to be motivated, and to feel, in certain ways about certain things, namely, to have an aversion to breaking a promise, to feel guilty about doing so,” that is, to act according to utility-maximizing rules of conduct. Implicit in this approach is the Aristotelian-Humean belief that the explanation for conduct reflecting character is internal to the actor and his desires, even if not fully within his control. On Brandt’s view, “a defect

---

<sup>54</sup> Christine M. Korsgaard, *THE SOURCES OF NORMATIVITY* 50 (Cambridge 1996).

<sup>55</sup> On rule utilitarian thought, see John Rawls, *Rule Utilitarianism*, in *PHILOSOPHICAL PERSPECTIVES ON PUNISHMENT* 82 (ed. by Gertrude Ezorsky State University of N. Y. Press 1972) (originally published as *Two Concepts of Rules*, 44 *PHILOS. REV.* 3 (1955) (distinguishing between justifications for the practice of punishment generally, and particular applications, and defending utilitarianism as an explanation for moral judgments thereunder); John Rawls, *Legal Obligation and the Duty of Fair Play*, in *LAW AND PHILOSOPHY: A SYMPOSIUM* 3, 9-10 (N.Y.U. Press 1964) (defining the duty of fair play as a scheme in which everyone benefits by mutual social cooperation and in which everyone or nearly everyone accepts certain restrictions on liberty to enjoy the benefit which is, in a sense, free, that is, the situation is unstable to the extent that one person knows that he can fail to cooperate but still enjoy the benefit).

<sup>56</sup> Richard B. Brandt, *A Utilitarian Theory of Excuses*, 78 *PHILOS. REV.* 337, 354 (1969) (emphasis in the original). Although I believe that the debate between character and choice theorists is substantially overinflated, see, e.g., Duff, *Choice, Character, and Criminal Liability*, *supra* note # –, I am assuming that character is a relevant factor in the search for blame, especially when the issue of excusing misconduct arises. See, e.g., M. P. C., cmt. to §2.04, at 275 (noting that the excuse of mistake of law applies when the “act charged is consistent with the entire law-abidingness of the actor”); *id.* §210.3, at 55 (stating that provocation acknowledges that “one who kills in response to certain provoking events should be regarded as demonstrating a significantly different character deficiency than one who kills in their absence”).

<sup>57</sup> Brandt, *supra* note # –, at 354 (emphasis added).

<sup>58</sup> *Id.*

*of character is, or includes, a defect of motivation.*<sup>59</sup> Presumably this means that the actor *could* have learned appropriate conduct but failed to do so because he lacked the appropriate motivational reinforcement. (In fact, whether any particular individual *could* be trained, in the sense that he has adequate capacities-in-fact for cognition and volitional control and has had the opportunity to learn same, is not an issue on which Brandt opines.<sup>60</sup> I suspect, though, that the dual requirements of capacity-in-fact do not present him with a forbidding hurdle.<sup>61</sup>)

Working in a similar vein, Norvin Richards dismisses the importance of constitutive luck on our assessments of human conduct.<sup>62</sup> (Constitutive luck describes the conventional understanding that our “inclinations, capacities, and temperament” – those relatively permanent qualities that some distance toward composing us – are to some immeasurable extent outside our control.<sup>63</sup>) Along with Brandt, Richards argues that such factors would count as exculpating, but only “if one’s character is *to no extent* one’s own artefact. But if the individual makes *any contribution whatever* to the sort of person he is, that

---

<sup>59</sup> Brandt, *supra* note # –, at 355 (emphasis added). Brandt’s survey of excusing conditions is fairly typical except insofar as he presses the motivational button: “All the considerations traditionally recognized as exculpating excuses are ones evidencing adequate motivation – or at least showing absence of evidence of inadequate motivation.” *Id.* at 357. Notably absent is serious consideration of volitional deficits or how they are acquired. In fact, is understanding of a motivational defect begs the question how it came about: “For part of what it is to *blame* a person is to impute a *defect* to him – and a defect is a dispositional feature of him, an incapacity or deleterious tendency.” *Id.* at 385 (emphasis in the original).

<sup>60</sup> Part of my critique of Brandt’s view is a reaction to its neurobiological naivete. No serious neuroscience scholar doubts that genetic endowment, for example, plays a substantial role in the construction of personality. *See, e.g.* David Wasserman and Robert Wachbroit, *Introduction: Methods, Meaning, and Morals*, in *GENETICS AND CRIMINAL BEHAVIOR* 1, 12-13 (ed. by David Wasserman and Robert Wachbroit: Cambridge 2001); Kenneth Taylor, *On the Explanatory Limits of Behavioral Genetics*, in *GENETICS AND CRIMINAL BEHAVIOR*, *supra* at 117, 125-26 (summarizing data that suggest that some criminality reflects defects in an individual’s “basic cognitive or affective architecture, citing B. Maughan, *Childhood Precursors of Aggressive Offending in Personality-Disordered Adults* in S. Hodgins, *MENTAL DISORDER AND CRIME* (1993) (noting that while only 7% of the general population suffers from antisocial personality disorder, 45% of a population of convicted violent felons so suffer)); M. Virkkunen and M. Linnoila, *Serotonin in Personality Disorders with Habitual Violence and Impulsivity* in Hodgins, *supra* (linking habitual aggressive behavior with reduced serotonin levels); and J. Volavka, D. Martell, and A. Convit, *Psychobiology of the Violent Offender*, 7 *J. FORENSIC SCI.* 237 (1992) (suggesting that reduced serotonin follows from a genetic mutation)). Yet it is precisely these developments that our jurisprudence is just now beginning to take any cognizance of. *See* note – *infra* (sources cited).

<sup>61</sup> *Cf.* Brandt, *supra* note # –, at 351-52 (stating that excuses entail societal costs insofar as the wrongdoer or potential wrongdoer believes that the clever lawyer will help him use the excuse to avoid incarceration).

<sup>62</sup> Norvin Richards, *Luck and Desert*, 95 *MIND* 198 (1986).

<sup>63</sup> Thomas Nagel, *Moral Luck*, in *MORTAL QUESTIONS* 38 (Cambridge 1979) (originally published in 50 *PROCEEDINGS ARISTOTELIAN SOCIETY* (Supp. 1976)).

contribution can be the basis of for his deserving . . . blame for what he does.”<sup>64</sup> Although he concedes that bad luck plays some role in determining one’s character, he opines that we can also make wrong choices, and it is the “*wrong part*” of human conduct that he want to take “seriously.”<sup>65</sup>

Richards’ implicit proposition – that excuses are proper only if our character is “*to no extent*” of our own making and are inappropriate, therefore, if one’s character makes *any contribution whatever* to our own authorship – raises at least three separate but related questions that intrude directly upon our judgments about blameworthiness and criminality. The questions implicate (a) the ontological proposition that we are to some extent the authors of our own lives; and (b) epistemologically-informed metaethical/ethical questions of (i) whether we possess the tools to *recognize* correctly those features over which we do and do not have authorial control, and (ii) whether we can, therefore, accurately *attribute* particular instances of misconduct to those circumstances over which we have such control and, thereafter, impose justifiable penalties on account of that over which we do exert control.<sup>66</sup> Simply to clarify, much less disentangle the polycentric issues embedded within the phrases “*to no extent*” or “*makes any contribution whatever*” illustrates the difficulties with this position. For example, while I readily accept the “authorial” premise (a) as given, at least when the agent has sufficient cognitive capacity-in-fact and opportunities to develop a mature conscience, I cannot embrace propositions (b) or (c) because recognizing such features and genuinely attributing responsibility are far more difficult than the Brandt-Richards position allows, a point I take up shortly.

---

<sup>64</sup> Richards, *supra* note # –, at 202 (emphasis in the original).

<sup>65</sup> *Id.* (emphasis added). George Fletcher’s approach to excuses indicates similar views, although his final account is unclear. Stated in summary form, Fletcher contends that because punishment is justified only if measured in terms of desert, where desert is a measure of character, we are required to make a judgment about the actor’s character whenever we adjudicate the propriety of excusing certain conduct. George Fletcher, RETHINKING CRIMINAL LAW §10.3.2, at 800 (1978). The points of view taken here are akin to, if not instantiations of, what Professor Susan Wolf describes as the “deep-seated view” of the free self, according to which we are responsible for our actions not only because they are intentional, but also because they are necessary reflections of a will that expresses our deep seated character traits. Susan Wolf, *Sanity and the Metaphysics of Responsibility*, in RESPONSIBILITY, CHARACTER, AND THE EMOTIONS, *supra* note # –, at 46, 47-50 (discussing the work of Harry Frankfurt, Gary Watson and Charles Taylor). As Wolf points out, these positions, like those discussed above, leave open the question: “Who, or what, is responsible for this deeper self?” *Id.* at 51 (concluding that the ability to step back and contemplate our own actions, the standard reply, is insufficient).

<sup>66</sup> Perhaps Professor Richards meant no such thing. Perhaps he has or would recommend issues related to cub-pathological capacity deficits be dealt with at sentencing. See, e.g., §5K.2.13, SENTENCING GUIDELINES MANUAL (permitting a departure for “significantly reduced mental capacity,” provided, *inter alia*, that the offense charged is non-violent). If that is the case, then, we must (at the very least) reappraise the propriety of the assumptions in our fairly non-discretionary Sentencing Guidelines. See, e.g., Thomas W. Hutchison, David Yellen, Deborah Young, and Matthew R. Kipp, FEDERAL SENTENCING LAW AND PRACTICE, ch. 5, pt. H (St. Paul: West, 1999) (excluding from consideration the conditions in which wrongdoers, through accidents of birth alone, are raised and receive – or do not receive – a moral education). At worst, we must ask if he hasn’t buried issues related to the presumption that we can genuinely ascribe moral responsibility more deeply and less rationally than it is presently.

All of the basic questions raised by this view of character rest on the problematic assumption that all those wrongdoers whom the law deems sane and competent – BK, Yates, Harris, Dahmer and the like – have both control over their character and that character in general is relatively stable over time.<sup>67</sup> Although the latter point is contestable,<sup>68</sup> one can suppose for now that one’s moral development goes through stages and becomes more-or-less fixed at some point in time with respect to at least some areas of conduct. Even if this is correct, however, the point overlooks or dismisses as unimportant an incontrovertible fact: Who we are at *t*, the moment of wrongdoing on which we focus for rendering moral and legal accounts, is a function of all the antecedent influences whose impact on the decision made at *t* are certain, albeit not fully discernible in the aggregate.<sup>69</sup> Nor will it do to claim, as many do, that because determinism is

---

<sup>67</sup> Cf. Elkhonon Goldberg, *THE EXECUTIVE BRAIN: FRONTAL LOBES AND THE CIVILIZED MIND* 150 (Oxford, 2001) (suggesting that data from neuroscience require a “new legal construct of ‘inability to guide one’s behavior despite the availability of requisite knowledge’ may be needed to capture the peculiar relationship between frontal lobe dysfunction and the potential for criminal behavior”). See *id.* at 108-09 (describing a phenomenon characterized by the “non-pathological diminution of the ability to form a theory of the mind,” i.e., the inability to sense how others are reacting internally to us as reflecting “normal variability” in frontal lobe functioning).

<sup>68</sup> See John M. Darley and Thomas P. Shultz, *Moral Rules: Their Content and Acquisition*, 41 *ANNU. REV. PSYCHOL.* 525, 541-45 (1990) (discussing research on how children learn moral rules); Yoram Shachar, *The Fortuitous Gap in Law and Morality*, 6 *CRIM. JUST. ETHICS* 12, 24 (1987) (surveying Piaget’s work among others and speculating that “human morality, once autonomously developed and formed, remains so in *some* areas of conduct, while other authorities replace parental authority in other areas and create in the human mind the same subservient nonautonomous opinion which is a mere reflection of external manifestations of attitude by such authorities”). As Daniel Goleman, *EMOTIONAL INTELLIGENCE* xiii (1995) notes, “the neurological data suggest [the existence of] a window of opportunity for shaping our children’s emotional habits.”

Note, in this regard, the increasingly questionable but persistent unstated assumption of adequate parental or adult authority. On the other end of the developmental chain, there is some evidence that genetic engineering may be employed to improve some aspects of personality that one does not like. See Martha Nussbaum, *Brave Good World*, *THE NEW REPUBLIC*, Dec. 4, 2000, at 38 (reviewing Allen Buchanan, Dan W. Brock, Norman Daniels, and Daniel Wikler, *FROM CHANCE TO CHOICE: GENETICS AND JUSTICE* (Cambridge U. Press 2000)).

<sup>69</sup> Bernard Williams makes this point about the instability of character judgments made at, or in reference to, a fixed point in time, in terms of the criteria needed to make such a determination:

[W]hat one does and the sort of life one leads condition one’s later desires and judgments. The standpoint of that retrospective judge who will be my later self will be the product of my earlier choices. So there is no set of preferences both fixed and relevant, relative to which the various fillings of my life-space can be compared. If the fillings are to be evaluated by reference to what I variously . . . want, the relevant preferences are not fixed, while if they are to be evaluated by what I now . . . want, this will give a fixed set of preferences, but one that is not necessarily relevant.

Bernard Williams, *Moral Luck*, in *MORAL LUCK: PHILOSOPHICAL PAPERS, 1973-1980*, 29 (Cambridge 1981) (rejecting, generally, the Kantian hypothesis that good or bad will is “unconditioned” [or] . . . free from external contingency”).

universal and affects all of us, it can be dismissed.<sup>70</sup> For now, suffice it to say that “universal” is not the same as “uniform.” We are each affected uniquely: some for better, some for worse.<sup>71</sup> This brute social fact of existence challenges our conception of freedom because the conventional wisdom posits that free actions “should not be determined by antecedent conditions and should be fully explained only intentionally, in terms of justifying reasons and purposes.”<sup>72</sup> Yet antecedent events and conditions obviously *do* matter greatly, and everyone knows it; our well nurtured intuitions shout out this basic understanding routinely.<sup>73</sup> Even our discourse about character and personality betrays the incompleteness of our current understanding of blameworthiness, an incompleteness in its expressions and judgments and, ultimately, in its practical wisdom.

B. The abbreviated semantics of character – Perhaps it is not surprising that character theorists tend not to define character with great rigor.<sup>74</sup> It’s a wobbly term, the contours of which are just now giving way to serious study. So, like the late Justice Potter Stewart, they assume that they recognize the impact of character when they see it,<sup>75</sup> and they know it for certain when it produces an unwelcome outcome; that is to say, blamers tend to reason in a *post hoc ergo propter hoc* manner.<sup>76</sup> The task of attributing blame

---

<sup>70</sup> See, e.g., *Moral Excesses*, *supra* note # –, 75-6 McGeorge L. Rev. 63 (2000); Moore, *Causation and the Excuses*, *supra* note # –, at 1092 (arguing that “moral responsibility for an action should be ascribed to an actor even when that action was caused by factors over which he has no control”); Morse, *Moral Metaphysics*, *supra* note # –, at 886.

<sup>71</sup> This is a basic fact of our genes and neurobiology. See, e.g., Ridley, *supra* note # –; Damasio, *supra* note # –.

<sup>72</sup> Thomas Nagel, *THE VIEW FROM NOWHERE* 115 (New York: Oxford 1986). See also Thomas Nagel, *WHAT DOES IT ALL MEAN?* 51 (New York: Oxford 1987).

<sup>73</sup> Even our ordinary, everyday usage underscores this fact. We routinely talk to our friends and neighbors across fences – and without ever second-guessing ourselves (a human quality of mixed virtue) – about our genetic make-ups, including our likenesses to Mom or Dad (“like father, like son,” “the acorn never falls far from the tree,” and so on), to our sisters or brothers, and the variations nature has carved out among our siblings. On the importance of family selection studies, see, e.g., Robert Plomin, *NATURE AND NURTURE: AN INTRODUCTION TO HUMAN BEHAVIORAL GENETICS* 38-41 (Pacific Grove, Cal: Brooks/Cole Publ. Co. 1990).

<sup>74</sup> See, e.g., Richards, *supra* note # –, at 200 (defining character as a person’s “set of preferences, in so far as these are relatively stable over time”).

<sup>75</sup> *Jacobellis v. Ohio*, 378 U. S. 184, 197 (1964) (declining to define the term “obscenity,” but noting that “I know it when I see it.”) (Stewart, J., concurring).

<sup>76</sup> Brandt, for example, fully discusses character as a factor in making moral assessments, but gives short shrift to its definition by declaring that it reflects “relatively enduring response tendenc[ies]” revealed in contexts rife with moral implications. Richard Brandt, *ETHICAL THEORY* 466 (Prentice Hall 1959). Note that this “definition” moves backwards, from effect to cause; it follows from a purported effect that assumes character as its cause. Richards defines character simply as one’s “set of preferences, in so far as these are relatively stable over time.” Richards, *supra* note # –, at 200. For a generally sympathetic (but not fully convinced) view of Brandt’s theory, see

based on misconduct that reflects the bad content of one's character is a tricky business, as likely to reveal those cognitive biases of the observer as it is the bad character of the actor. In this section, I look first at the complex nature of describing fully one's character. Thereafter I report on the social psychology of bias that accompanies attributions of blame.

The search for an individual's bad character commonly appears in our folk psychological usages as familiar shorthand expressions that serve as proxies for rich descriptions of complex human behavior. This idiomatic simplicity meets *our* immediate needs to release anger, define blame, evoke shame, and move on. We use naive terms to describe one's "character" – such as sweet or shy or considerate or clumsy or surly or vicious or bad – when we want to say something about the observed actor's supposed intentions sufficiently general so that we then can use our own experiences and imaginations to fill in the epistemic gaps in our attributions of character. That is, we use familiar metaphors in lieu of fully informed psychological explanations. These semantic shortcuts serve us well enough in everyday conversation, but they well fall short as accurate descriptions and thus fail to meet our normative goal as a flourishing, compassionate, justice-seeking community.

The semantic case is made in part (and reluctantly) by Adam Morton, who notes that the function of conceptions of character in psychological explanation is "to block, qualify, or emphasize the connection between beliefs and desires that an explanation appeals to and the actions it explains."<sup>77</sup> We use general descriptions of character traits that reveal *some* relatively common patterns of observation relevant to the peculiarities of different incidents. In this partially-informed way, our common metaphors permit the attributor to avoid describing character in behavioral terms. The simple explanation for this process is that fully-described behavioral states always seem "too complex for a quick explanation."<sup>78</sup> So terms like "brutal," "vicious," "angry," "short-tempered," and "bad" generally truncate discussions about the compound states that actually generate behavior. Instead, we find quick familiar (visible) connections between complex internal states and the actions we observe. We use the distancing term "character" in lieu of dealing with the "person" of the accused.

In a general though still incomplete way, we follow this condensed approach because a genuine, "ultimate perfect character term," *T*, for describing one's character would go something like this:

A set of states *S* and the term [*T*] would jointly hold only when one of a

---

Andrew E. Lelling, *A Psychological Critique of Character-Based Theories of Criminal Excuse*, 49 SYRACUSE L. REV. 35, 47-59 (1998).

<sup>77</sup> Most of the discussion in this section is based on Adam Morton, *Character and the Emotions*, in EXPLAINING EMOTIONS 153, 155 (ed. by Amélie O. Rorty; Cal 1980).

<sup>78</sup> *Id.*

defining range of actions was performed, and given an action there would be a set of combinations of *S* and [*T*] which account for it.<sup>79</sup>

Something close to *T*, a more accurate (if not perfect) character description, requires that we describe these *states* with care, as the products of unique world experiences that emerge from something more than the details revealed by a snapshot taken at *t*. A more fully descriptive term or phrase would encompass a substantial record of the various states that produced this unique person: the *events, conditions* (including not least genetic endowment, significant trauma and so on) and, as importantly, the *opportunities* to trigger appropriate responsive “states.” Acting together, these are the circumstances that generate one’s “character,” and it is critical to note that the environment – each individual’s *own* peculiar environment – triggers genetic change and neuronal growth and death in a process of constant self-origination.<sup>80</sup>

A more fully reported *S*, therefore, must include both analytically-appraised intentional behaviors and their related neurobiological adaptive processes.<sup>81</sup> An *S* that incorporates the antecedent circumstances affecting character development over time – personhood – would reveal a complicated, interconnected network of traits, but it is a process over which we should seek maximum understanding because doing so will yield policies and decisions that embrace genuine compassion and move toward greater safety over time.<sup>82</sup> The *T* that more fully describes an individual’s intentions no more complex

---

<sup>79</sup> *Id.* at 156.

<sup>80</sup> The point here is that whereas some genes guarantee that the overwhelming majority of individuals are born with all and only the right parts in the right places, they have other functions as well. “The function of many genes is . . . to switch other genes on or off. And the susceptibility of a gene to be switched on or off depends on the sensitivity of its promoters,” that is, on a species of genetic material that facilitates the production of proteins when other genetic materials (“transcription factors”) attach themselves. Ridley, *NATURE VIA NURTURE*, *supra* note # –, 31. What causes such genes to switch other genes on or off? Put simply, the environment, defined broadly in terms of the unique, non-genetic experiences each individual encounters. An individual’s *S* (or motivational set) includes at least one’s “desires, evaluations, attitudes, projects, and so on.” Williams, *Internal Reasons*, *supra* note # -- at 35. Thus, describing the composition of a *state* is itself a complex task. *Cf.* David C. Rowe and Kristen C. Jacobsen, *In the Mainstream: Research in Behavioral Genetics* 23-4, in Ronald A. Carson and Mark A. Rothman, *BEHAVIORAL GENETICS: THE CLASH OF CULTURE AND BIOLOGY* (Johns Hopkins 1999) (describing *state dependence* as part of a developmental model according to which “past behavior affects future behavior,” such that the commission of a crime increases the likelihood of future criminality). I discuss this in considerably more detail in *On the Potential of Neuroscience: A Comment on Greene and Cohen’s For the Law, Neuroscience Changes Nothing and Everything*, ----.

<sup>81</sup> See Allan Gibbard, *Genetic Plans, Genetic Differences, and Violence: Some Chief Possibilities*, in Wasserman and Wachbroit, *GENETICS AND CRIMINAL BEHAVIOR*, *supra* note # –, at 177-78 (noting the theoretical deficiencies of folk psychological explanations of causal relations).

<sup>82</sup> It is true that if the jurisprudence of excuses were taken to its logical conclusion, it reasonably could produce an environment in which all intelligible rationales for misconduct would produce an excuse, and that this environment is normatively unacceptable, at least without changing that jurisprudence. See Michael S. Moore, *Causation and the Excuses*, 73 CAL. L. REV. 1091, 1094 (1985) (“If one accepts the determinist thesis that all events,

than are the basic ideas of fairness and justice. Again, what is remarkable (even mind-boggling) about this debate is that *no one* seriously doubts the influences of such antecedent events and conditions on the development of character as a matter of fact and common intuition; it is only as a matter of our penal law and its expectations that we disregard these insights. Taking cognizance of these facts in our criminal justice system demands a good deal of honesty than we now exhibit and much more good will than our politics currently allow.

Instead of using more precise compound descriptions, which requires an expanded narrative of life – a sensitive inquiry into the adaptations that the actor’s environment required for survival – we substitute “fixed patterns of explanation in terms of beliefs, desires, and other well-behaved states,” which move us in this direction but stop well short, thereby encouraging us freely to fill in the epistemological chasms in light of patterns of character attributions we have already assimilated.<sup>83</sup> What we require, if we hope to find and reinforce the wrongdoers’ redemptive features, is the ability “to capture in one’s ascriptions of character the reasons why this particular person, at this time, is moved by these motives.”<sup>84</sup> Sadly, we don’t, in general, follow this course in our substantive law. The complexities always seem too great. Instead, Morton points out, “we invent a labyrinth of character terms interacting in their effects and stratified in complicated ways, to capture more and more of the particularity of people.”<sup>85</sup>

Within the highly cropped terms we use to describe character are patterns of notable conduct that we attribute to the bad character of wrongdoers in light of *our* collective

---

including all human behavior, are caused, and if one believes that causation excuses, then one must believe that moral responsibility is an illusion on which liability to criminal punishment cannot be built.”). So our jurisprudence concerning punishment and incarceration needs to be debated once again. “Punishment” and “incarceration” can be achieved more humanely with noun substitutes, like “treatment,” where both normatively and descriptively appropriate, in lieu of the current punishment default, and “incapacitation” when the risk of future dangerousness so requires.

This would be a daunting task if the prosecution were required to bear the burden of proof on “treatability,” and that failing to carry that burden produced a not “guilty” verdict. “Treatability,” for example, asks whether the defendant is more likely to achieve a more law-abiding status as a person undergoing extensive counseling and education than he is to emerge (if ever) as a hardened career criminal. We should strive to measure such tendencies with the best available medical and psychological standards and rules. I address the neurobiological basis of the measurement problem at length in *Blaming Angry Men*, *supra*,, mss in the author’s possession.

<sup>83</sup> Morton, *supra* note # –, at 156. On the systemic biases we bring to the attribution of blame, *see infra* text accompanying notes —.

<sup>84</sup> Morton, *supra* note # –, at 156. One apparent exception to this point is the nascent defense of diminished capacity. *See, e.g.*, *United States v. Brawner*, 471 F.2d 969 (D.C. Cir. 1972) (approving the admissibility of evidence of reduced mental capacity for the purpose of showing that the defendant was incapable of forming the requisite intent). One wonders, however, why a separate doctrine is required to address what is primarily an evidentiary issue. *See* M. P. C. § 4.02, cmt. at 219.

<sup>85</sup> Morton, *supra* note # –, at 156.



aspirations, which may or may not be accurate in light of *their* relative capacities-in-fact. In everyday matters, it seems “implausible and inaccurate to explain an inappropriate attitude by attributing a belief or desire that would rationalize it, because the apparently anomalous emotion is embedded in a system of . . . inappropriate attitudes or false beliefs.”<sup>86</sup> To assess these inappropriate (and often criminal) systemic attitudes with greater precision enjoins us to trace an individual’s causal history, and an apparent problem with this is prescription is determining the biographical starting point. For a full explanation one has to go all the way back “to explain the prepropositional but intentional habits of salience, organization, and interpretation. It is these which, through later intervening beliefs and attitude – many of them false and inappropriate – explain the conservation of emotions,”<sup>87</sup> that is, the persistence of evidently disadvantageous emotional (and thus behavioral) states.<sup>88</sup> The Greeks labeled some of these phenomena “*akrasia*;<sup>89</sup> we might call them “maladaptations.”<sup>90</sup>

C. The problematics of blame attribution – Unfortunately, the problem that is the everyday use of incomplete semantic descriptions of character in criminal law only begins to state the deeper problems with blaming. In order to factor in all the variables that would produce a better external understanding of one’s character, we must first cross another hurdle before we could accurately assess blameworthiness: namely, overcoming our own biases in the construction of blame attribution, a difficulty which, if overcome, could produce a reconfiguration of the burden of proof on issues such as basic competency.<sup>91</sup> Cognitive social psychologists use the phrase “Fundamental Attribution Error” (FAE) to describe a well studied phenomenon that reflects the existence of common biases in our attributions of wrongdoing: There is a significant likelihood that in any instantiation of our ordinary assessments of wrongdoing we over-attribute negative outcomes to character and disposition rather than the situations in which actors find themselves. In this subsection,

---

<sup>86</sup> Amélie O. Rorty, *Explaining Emotions*, in EXPLAINING EMOTIONS, *supra* note # –, at 103, 104.

<sup>87</sup> *Id.* at 103 (seeking to explain the “unexpected conservation of emotions,” that is, why people fail to change their emotional commitments when circumstances seem to warrant it).

<sup>88</sup> For a description of how such an evaluation might be made, *see* Gibbard, *supra* note # –, at 188-92.

<sup>89</sup> *See, e.g.*, Donald Davidson, *How is Weakness of the Will Possible*, in ESSAYS ON ACTIONS AND EVENTS 22, 31-3 (concluding that Aristotle provided no fully satisfactory explanation for “incontinent actions.”).

<sup>90</sup> Some of the most creative work in this area has been done by Owen D. Jones. *See, e.g.*, *Time-Shifted Rationality and the Law of Law's Leverage: Behavioral Economics Meets Behavioral Biology*, 95 NW. L. REV. 1141 (2001); *On the Nature of Norms: Biology, Morality and the Disruption of Order*, 98 MICH. L. REV. 2072 (2000); *Evolutionary Analysis in Law: An Introduction and Application to Child Abuse*, 75 N. C. L. REV. 1117, 1200-04 (1997).

<sup>91</sup> Just as we use second order “avoidance” techniques to ascribe blame when doing so appears to be consistent with our world views, *see, e.g.*, Lawrence M. Solan, *Cognitive Foundations of the impulse to Blame*, 68 BROOKLYN L. REV. 1003, 1004-05 (2003), we might need to rethink our default tests of competency and modifications to the burden of proof to counteract these tendencies.

I briefly explain Attribution Theory (AT) and describe the biases it produces, FAE. In the following subsection, I suggest how these biases operate pervasively to produce counter-productive results in the domain of punishment.

AT describes how individuals explain causation, the *whys* of human conduct that we routinely question: Why did this human act occur? When an actor performs an unwelcome event,  $\Phi$ , the observer often wants to know whether  $\Phi$ ing occurred because of something within the actor – that is, was it “dispositional” – or whether it occurred because of something within the actor’s environment – that is, was it “situational?”<sup>92</sup> AT evaluates the assessment procedure of ordinary individuals, jurors and judges, for example. Fritz Heider, the Gestalt psychologist who developed the theory, described his approach to AT as an example of “naive psychology” because it addresses how the person-in-the-street explains observed behavior.<sup>93</sup> Heider maintained that individuals account for the conduct of others based on their (the observer’s) own beliefs, whether those beliefs are valid or not.<sup>94</sup> The “alert perceiver” makes inferences about another person’s intentions and dispositions based on his or her observations; the inferences drawn are deemed to be “correspondent when the behavior and the disposition can be assigned similar labels.”<sup>95</sup> For example, the observer might conclude that because *X* did an act that she, the observer, views as mean-spirited (or kind), *X* is viewed as a mean-spirited (or kind) person. Observers seek correspondence. But is the observer’s naive psychological assessment of character (even usually) correct?

The answer is resounding and unsettling: sometimes. To begin with, attributing misconduct to an individual’s disposition or character raises the familiar problem of free will: The observer assumes that the actor being blamed enjoyed freedom of choice – he could have done otherwise.<sup>96</sup> The more difficult the action observed, moreover, the more

---

<sup>92</sup> These phenomena are sometimes referred to as “internal attribution,” which examines the causative forces internal to the actor that contribute to *X* doing  $\Phi$ ; and “external attribution,” which examines the causative forces external to the actor that contribute to *X* doing  $\Phi$ . Taken to their extremes, the former assigns only internal motivators to *X* doing  $\Phi$ , and the latter assigns only external motivators to *X* doing  $\Phi$ . Attribution Theory (I): Fundamentals, at [www.psyc.abdn.ac.uk/homedir/amilne/Social](http://www.psyc.abdn.ac.uk/homedir/amilne/Social).

<sup>93</sup> Attribution Theory, [http://hsc.uusf.edu/~kmbrown/Attribution\\_Theory-overview.htm](http://hsc.uusf.edu/~kmbrown/Attribution_Theory-overview.htm).

<sup>94</sup> For a still useful summary of Heider’s early work in Gestalt theory, see Morton Deutsch and Robert M. Krauss, THEORIES IN SOCIAL PSYCHOLOGY 29-36 (Basic Book 1965). According to Heider, “the attribution of personal responsibility involves a decision as to which of several conditions of action – the intentions of the person, personal power, or environmental factors – is to be given primary weight for the actual outcome. In general, the more the environmental factors are thought to influence the action, the less the person is held responsible for an action with which he is connected.” *Id.* at 32.

<sup>95</sup> Attribution Theory, *supra* note # – (attributing the phrase “alert perceiver” to disciples of Heider, E. E. Jones and K. E. Davis).

<sup>96</sup> There are at least two aspects to “free will,” although neither is sufficient to explain free will. The first is *freedom of action*, the second is *freedom of choice*. Freedom of action speaks to capacity: An actor could lack free will simply because he or she lacks a certain capacity to act freely. For example, a person with profound cerebral

likely the observer is to attribute it to free choice. For example, although it is probably not easy to split someone's skull in half with a hammer, an observer provided with little or no information before deciding whether a severe head injury resulted from accidental or intentional conduct is more likely to attribute the injury to an actor's bad character (or disposition) than to the situation.<sup>97</sup> This over-estimation is a constant phenomenon, and so there is a consistent "fundamental attribution error" running through the process: We tend to underestimate situational influences, and overestimate dispositional influences, on behavior.<sup>98</sup> Additionally, FAE is more prominent when we judge others than it is when we judge ourselves.<sup>99</sup>

When we attribute blame and responsibility to an individual for  $\Phi$ , which invariably results from a number of substantial causative factors, we tend to "cite the cause for which an actor is most blameworthy, that is, we search for the cause containing the most indelible stain" on the actor's character.<sup>100</sup> It is certainly not clear why this tendency to point and attribute fault to character has evolved and with what rewards and anguish, but it is an omni-present feature of the blaming practice.<sup>101</sup> Nor is it entirely clear what to do about this tendency beyond recognition. That something must be done, however, follows from the obvious fact that this systemic tendency does not reflect our better, more compassionate natures.

Professor Mark Alicke demonstrated this dispositional preference in an artfully designed study that asked subjects to review one of several experimental narratives that ended with the exact same bad outcome. Alicke asked subjects (observers) to evaluate the identical conduct of four individuals. The motives of the individuals involved in the narratives varied from favorable to benign to culpable; their motives, however, were

---

palsy lacks the capacity to undertake certain freely chosen activities, like playing basketball or serving as a stenographer. Freedom of choice, in contrast, speaks to opportunity. A person may have the evident capacity to choose, but lack the opportunity at the relevant time. See John Martin Fischer, *Introduction: Responsibility and Freedom*, in John Martin Fischer, ed., *MORAL RESPONSIBILITY* 9, 10-15 (Cornell 1986).

<sup>97</sup> Cf. Em Griffin, *A FIRST LOOK AT COMMUNICATION THEORY* (McGraw-hill, 2d ed. 1994) available at [www.afirstlook.com/archive/attribut.cfm?source=archther](http://www.afirstlook.com/archive/attribut.cfm?source=archther).

<sup>98</sup> *Id.* (quoting Lee Ross, *The Intuitive Psychologist and His Short-Comings: Distortions in the Attribution Process*, 10 *ADVANCES IN EXPERIMENTAL PSYCHOL.* 184 (1977) (page cite omitted)).

<sup>99</sup> See, e.g., Focus on Attribution Theory, at [www.ua1vm.ua.edu/~droskos/Attribution%20theory/sld002.htm](http://www.ua1vm.ua.edu/~droskos/Attribution%20theory/sld002.htm).

<sup>100</sup> Mark D. Alicke, *Culpable Causation*, 63 *J. PERSONALITY & SOC. PSYCHOL.* 368, 369 (1992) (describing the "stain theory" of Joel Feinberg, *Sua Culpa*, in Feinberg (ed.) *DOING AND DESERVING: ESSAYS IN THE THEORY OF RESPONSIBILITY* 187 (Princeton 1970)).

<sup>101</sup> It could be that we use blame to channel retributive urges that many people cannot control. Nonetheless we understand intuitively that at the soul of retribution is anger and the potentially vicious conduct it can generate. See *Some Thoughts on the Aesthetics of Retribution*, note – *supra*.

relevant only to the occasion of the harm, and not to its material causation. Despite the absence of this causal relevancy, the subject-observers significantly attributed fault to the actors whose motives for action were perceived as more blameworthy (less worthwhile) than to those actors whose motives were neutral or favorable. To reiterate, the observers made the negative character attributions notwithstanding the fact that motivation was causally unrelated to the unwelcome outcomes.<sup>102</sup>

There are several ways to explain these and similar findings, and none is necessarily exclusive of any (or all) others for they can be synthesized based on existing research data. Fincham and Schultz, for example, advance an “entailment model” to explain FAE. According to their basic rationale, the tendency to blame (in general) occurs whenever observers discover some causal link between the actor and the harm inflicted. Importantly, *that* causal linkage is sufficient to impose retributive punishment on these actors, because the observer’s attribution of causal control establishes not only the necessary evidential link between the actor and the harm inflicted; it also serves in the process of justifying punishment.<sup>103</sup>

Weiner moves the explanation further along with his identification of a “stage theory,” according to which our assessments of responsibility generally move in stages from “an evaluation of personal and impersonal causation (based on an analysis of intent) and considers the extent to which the actor controlled the harmful outcomes and [then to] whether mitigating circumstances existed.”<sup>104</sup> Weiner provides empirical support for a basic

---

<sup>102</sup> For example, in one experimental setting, a driver (D1) exceeding the speed limit (40 mph in a 30 mph zone) hit a car traveling in the opposite direction; the “other” driver (D2) suffered multiple cuts, a broken collar bone and a fractured arm. Observers were then told that the accident occurred in one of four contexts: as D1 (a) was rushing home to celebrate his anniversary; (b) was rushing home to hide a vial of cocaine he had left in the open at his parent’s home; (c) tried to stop but hit an unanticipated oil spill on the road; or (d) tried to avoid D2 who had run through a stop sign. The study showed that causation was attributed to D1 as a function of his motive, and that D1 was more likely to be held responsible (and pay enhanced damages) depending on the negative circumstances in which the accident occurred. As Alicke noted, “in the socially undesirable motive condition [the desire to hide cocaine], the focal actor [D1] was . . . cited more frequently as a cause than any other factor, even when [D2] was clearly negligent in running through the stop sign.” Alicke, *Culpable Causation*, *supra*, note # –, at 369-70. For similar findings see F. D. Fincham and T.R. Shultz, *Attribution of Responsibility: From Man the Scientist to Man as Lawyer* in 13 *ADVANCES IN EXPERIMENTAL SOCIAL PSYCHOLOGY* 81 (L. Berkowitz, ed., 1981).

<sup>103</sup> Mark D. Alicke, *Culpable Control and the Psychology of Blame*, 126 *PSYCHOL. BULL.* 556, 557 (2000) (describing Fincham and Shultz, *Attribution of Responsibility: From Man the Scientist to Man as Lawyer supra* note # –) (hereafter, “*Culpable Control*”). Put otherwise, “a judgement of blameworthiness presupposes a judgement of causation.” Kathleen M. McGraw, *Conditions for Assigning Blame: The Impact of Necessity and Sufficiency*, 26 *BRIT. J. SOC. PSYCHOL.* 109, 114 (1987); accord John M. Darley and Thomas R. Schultz, *Moral Rules: Their Content and Acquisition*, 41 *ANNUAL. REV. PSYCHOL.* 525, 531 (1990). This does not mean, however, that judgments of causation, blame and punishment are identical; only that they are closely inter-related.

<sup>104</sup> Alicke, *Culpable Control*, *supra* note # –, at 557 (summarizing B. Weiner, *JUDGMENTS OF RESPONSIBILITY: A FOUNDATION FOR A THEORY OF SOCIAL CONDUCT* (NY: Guilford Press 1995)).

assumption that is formally embedded (if not always actually honored<sup>105</sup>) in our law: a direct relationship between levels of criminal intent and desert. According to a third theory, described as a “sequential stage model,” observers typically move “into sequential stages of causal attribution, responsibility attribution, and finally blame,” where blame is broken down further into moral responsibility (knowing, voluntary action) modified (whenever appropriate) by an excuse or justification.<sup>106</sup>

Alicke’s contribution to this line of research is an extension of sequential stage theory called the “culpable control model.” This work adds two dimensions to the earlier research. First, it provides empirical support for the presence of a strong spontaneous affective component in the process of attributing blame. It also adds a “blame-validation” component, where evidential and motivational biases are deployed to validate the prior spontaneous ascriptions of blame. “Because cognitive shortcomings and motivational biases are endemic to blame,” Alicke writes, “a psychological treatment of this model requires a model in which personal expectations and emotional reactions are central components.”<sup>107</sup> This affective/emotional component nourishes the three inter-related elements of personal control assessment: “mental states, behaviors, and consequences,” where the mental element encompasses intentions, the behavioral elements include acts and culpable omissions, and the consequences include both immediate and extended attributions.<sup>108</sup>

Alicke’s model assumes that observers “deliberately and consciously” assess the links between (1) choice and conduct, (2) the actor’s apparent impact and the harmful outcome, and (3) his desire for and anticipation of the unwelcome consequences: all staples of criminal law. The novelty he contributes from the perspective of criminal law – and what puts him in the camp with those who study FAE – is his demonstration that individual observer judgments are influenced by “relatively unconscious, spontaneous evaluations” in the assessment of each of the three elements. “Spontaneous evaluations are affective

---

<sup>105</sup> For a useful work that ties FAE to the doctrinal anomalies of felony murder, to the attempt/completed crime asymmetry in punishment, and to its affect on different theories of punishment that defy the formal principle of grading, see Donald A. Dripps, *Fundamental Retribution Error: Criminal Justice and the Social Psychology of Blame*, 56 VAND. L. REV. 1383 (2003).

<sup>106</sup> Alicke, *Culpable Control*, *supra* note # –, at 557 (discussing the work of K. G. Shaver, *THE ATTRIBUTION OF BLAME; CAUSALITY, RESPONSIBILITY, AND BLAMEWORTHINESS* (NY: SPRINGER-VERLAG 1985)).

<sup>107</sup> Alicke, *Culpable Control*, *supra* note # –, at 557.

<sup>108</sup> *Id.* Griffin, *supra* note # –, describes a similar three-part process of attribution, including perception of action, judgment of intention, and attribution of intention. This approach to determinations of blameworthiness go back at least to Adam Smith, *THE THEORY OF MORAL SENTIMENTS* 133 (Prometheus Books 2000). See *infra* text accompanying notes –. I have addressed the “irregularity of sentiments” according to which we place outsized emphasis on (3) in terms of moral theory in *{Reference omitted.}*

[emotionally- induced] reactions to the harmful event and the people involved.”<sup>109</sup> They are driven by two sources: the information the observer possesses about the actor’s “intentions, behaviors, or the consequences they produce,” and “extra-evidential factors” such as the actor’s reputation and social status.<sup>110</sup>

What is of central concern here is the finding that our attribution biases are nearly universal and difficult to eradicate: The construction of blame follows from the perception of inappropriate motivation, even when motivation is literally irrelevant to the cause of harm. This tendency continues to plague us even when observers attempt to validate their initial spontaneous attributions of blame, which occurs when the stress of observation is lowered and the blame attributors (presumably) make their assessments more soberly; that is to say, when the rational analysis model would predict better judgment, observers still review the evidence of linkage from intention to behavior to consequences in the same biased manner.<sup>111</sup> Alicke’s work indicates that observers continually exaggerate the actor’s volitional control over unwelcome outcomes either “by lowering their evidential standards for blame or by seeking information to support their blame attribution bias.”<sup>112</sup>

This same phenomenon is reflected in our tendency to over-impute causal control to human agency, namely, FAE.<sup>113</sup> As a result of blame-validation processing, “observers are inclined to blame the actor or actors who arouse the most negative affect or whose behavior confirms unfavorable expectations.”<sup>114</sup> The upshot is that “expectations and

---

<sup>109</sup> Alicke, *Culpable Control*, *supra* note # –, at 558; accord Bargh and Chartrand, *supra* note # –, at 465 (“The idea that social perception is a largely automated psychological phenomenon is now widely accepted.”). See generally Russell H. Fazio, David M. Sanbonmatsu, Martha C. Powell and Frank R. Kardes, *On the Automatic Activation of Attitudes*, 50 J. PERSONALITY AND SOC. PSYCHOL. 229 (1986).

<sup>110</sup> Alicke, *Culpable Control*, *supra* note # –, at 558.

<sup>111</sup> *Id.* On rational analysis, see Robert Nozick, *THE NATURE OF RATIONALITY* 74 (Princeton, 1993) (stating that the “rational person will try to be aware of whether the reasons she has are a biased indicator of truth and will accordingly shape her estimate of the truth from the reasons she has. If she judges her reasons to be unrepresentative in some particular direction, she will correct for this.”).

<sup>112</sup> Alicke, *Culpable Control*, *supra* note # –, at 558. The phenomenon I’m elucidating here is an instantiation of cognitive dissonance. I have examined this well researched human reaction in a legal context in *The Third Best Choice: An Essay on Law and History*, 41 HASTINGS L.J. 537 (1990).

<sup>113</sup> The factors involved in assessing causal and volitional control vary as a function of the observer’s perception of whether (and to what extent) the actor’s conduct is seen as willful and purposive as opposed to accidental: the greater the perception of purposes, the greater the assessment of volitional control. Issues of personal control also vary as a function of assessments of individual capacity and situational constraints, roughly corresponding to cognitive capacity (diminished capacity, for example) and justifications and excuses. Alicke, *Culpable Control*, *supra* note # –, at 559-61.

<sup>114</sup> *Id.*

affective reactions [are] conflated with assessments of personal control.”<sup>115</sup> Because of these deep-seated, non-idiosyncratic reactions, we tend to alter our judgments of blameworthiness, generally assessing more blame on the putative wrongdoer than the context requires. This tendency to construct blame as reactions to perceived agential motivation and control is significant and significantly harmful.

D. Why FAE matters: FAEC (FAE + Costs) – The price we pay for a practice of judging that over-attributes wrongdoing to character, FAEC, is the subject of an important recent article by Professor Garth Williams. Williams notes that our current culture of blaming generates three counterproductive features that undermine the potential benefits our punishment regime might produce for our culture generally and wrongdoers specifically. Each of the three costs arises by virtue of the unwillingness of the blamer (or observer, which includes all of us individually and collectively, and not least jurors and judges) to confront the complexities that underlie our attributions of blame and responsibility.<sup>116</sup>

FAEC-1, the first negative effect attributable to over-emphasizing the individual’s character, is a tendency to demonize the actor. By failing to consider adequately the situational factors in which wrongdoing occurs, we come to hold unwarranted beliefs about the truth value of our assumptions about what an actor *could have done other than*  $\Phi$ ing, engaging in wrongdoing.<sup>117</sup> The result of this inclination threatens to focus all of our attention on the actor’s dispositions and intentions and, in so doing, yields a *practice* that simply does not serve the long term normative goal of making our neighborhoods safer places to live. By focusing on the actor’s character, observers (Williams call them “finger-pointers”) leave us with few opportunities to offer constructive help. (This might explain politicians’ disdain for reformative practices; the politician generally wants to present

---

<sup>115</sup> *Id.*

<sup>116</sup> Garth Williams, *Blame and Responsibility*, 6 ETHICAL THEORY AND PRACTICE 427 (2003). And, of course, the crabbed, binary understanding of capacity-in-fact often disables fact-finders even from considering (at least formally) the effects of substandard non-pathological cognitive and volitional deficiencies.

<sup>117</sup> *Id.* at 432; McGraw, *supra* note # –, at 115. Alicke also considers the effect of counterfactual reasoning and attributions of control, noting that if the harmful outcome would have occurred independent of the actor’s causal contribution, the perception of “effective causal control” is reduced but [as he shows] . . . not eliminated. Alicke, *Culpable Control*, *supra* note # –, at 561. This assessment appears to occur in two stages: First, the observer contemplates how the harmful outcome could have been avoided. Next, he accords causal priority to causal events that could have been avoided and uses that priority as a basis for attributing blame. See Neal J. Roese, *Counterfactual Thinking*, 121 PSYCHOL. REV. 133 (1997).

Whether we “come to” the use of these counterfactuals as a product of our inherent natures or borrow them from without is, from a Humean perspective, probably not relevant. See *supra* text accompany note – (declaring that the “sentiment of disapprobation, which arises in you, toward this action. . . lies in yourself, not in the object.”).

himself or herself as not just tough-on-crime but the opposite of “them.”<sup>118</sup>) The wrongdoer is too often reduced to a label: “just the kind of person who does  $\Phi$ ,” a designation that transforms the actor in the observer’s consciousness into someone who regularly acts inappropriately if not invariably criminally.<sup>119</sup> After describing the culprit in such hopeless terms, the alternatives left to the many finger-pointers among us concerning the actor’s future are narrow and gravitate toward “manipulation,” which could be broadly or narrowly conceived and is always prone to moving toward “outright coercion,” thus running the gamut from a slap on the wrist to death by gas or electrocution or lethal injection.<sup>120</sup>

Designating the wrongdoer as “*just the kind of person who*” shows the troubling metaphysical problem discussed previously: it threatens to render absurd our legal culture’s nearly irrebuttable presumption of free will in many instances. It is not strictly a violation of the law of non-contradiction to use this off-putting characterological description *and* simultaneously to avow that the wrongdoer commands free will, but it is at least puzzling and troubling.<sup>121</sup> Williams elaborates:

[I]t is somewhat odd to say that the culprit “*just does that sort of thing*” (dispositional attribution) and also freely, that is, *deliberately*, chose to act in that undesirable way. These are only consistent, and this is not at all coincidental, if we picture a more or less demonic agent, one who chooses wrong for wrong’s sake.<sup>122</sup>

Stated otherwise, the over-emphasis of dispositional blaming tends to transform wrongdoers into creatures like Melville’s Claggert, the personification of pure evil: The likes of Claggert, Melville wrote, are capable of

apprehending the good [they possess theoretical reasoning], but [are] powerless to *be* it [they lack full-throated practical reasoning]; a nature like [his] surcharged with energy as such natures almost invariably are, what recourse is left to it but to recoil upon itself and like the scorpion for which

---

<sup>118</sup> Even so-called “moderate” to “liberal” politicians embrace “get tough” policies when it comes to discussions of punishment in election politics. See, e.g., *The Cost of the Incarceration Boom: Correction Expenditures Rise, While Programs and Oversight Are Cut*, at <http://home.comcast.net/~lwva/Harshba.htm> (quoting former Massachusetts’ Governor William Weld stating that “prison should be like a tour through the circles of hell”).

<sup>119</sup> Williams, *supra* note # –, at 434.

<sup>120</sup> See *id.* at 434-35.

<sup>121</sup> The law of non-contradiction holds that believing *p* and *not-p* concurrently defines irrationality. See David Wiggins, *Towards a Reasonable Libertarianism*, in *ESSAYS ON FREEDOM OF ACTION* 48 (ed. by Ted Honderich (1973) (quoted in Bok, *supra* note # –, at 69).

<sup>122</sup> Williams, *supra* note # –, at 435 (first emphasis added).



the Creator alone is responsible, act out to the end the part allotted to it.<sup>123</sup>

This view runs counter to a significant feature of our moral and theological values – compassion.

This distancing, which is constitutive of a fundamental attribution error, produces FAEC-2, which Williams describes as “tunnel vision:” a narrowing of the perspective we bring to bear on our judgments about the actor. Over-attributing unwelcome outcomes to the actor’s character or disposition devolves into a refusal to take cognizance of the other’s perspective, which (in turn) undermines the wrongdoer’s ability to receive honest constructive criticism.<sup>124</sup> What is important is that blaming usually reflects an emotionally-laden reactive attitude that accents our negative view of the actor to the detriment of situational considerations, and its narrowing tendency makes clear a point this essay began with: Blaming is a complex social process in which all of us, actors and observers alike, are engaged and about which we must continuously ask questions. As such, it requires that we give to it its own commanding attention; we can’t focus only on what appears to lie behind  $\Phi$  because the process generates contradictions that do not serve us well. We cannot simply “lock ‘em up and throw away the keys,” as our overlapping retributive inclinations would often have us do.<sup>125</sup>

The affective/emotional aspect of FAE clarifies that rather than merely assessing and responding to the many  $\Phi$ s wrongdoing produces, we as observers and finger-pointers discharge our own emotions toward the actor; blaming reflects the internal mental states that we project onto an external reality. Another negative outcome of this practice is one that Joel Feinberg has described: We indelibly stain the wrongdoer by ratcheting up our disapproval, which rarely produces a beneficial encounter between blamer and actor.<sup>126</sup> At the same time and possibly more damning still, we suffer the tendency of reducing

---

<sup>123</sup> Herman Melville, *BILLY BUDD* 39 (Signet ed. 1998). Rawls, following Sedgwick, lists the following skills and attributes necessary to exercise full practical reasoning:

- (i) To adapt the most effective means to our ends.
- (ii) To acquire reasonable beliefs about our ends and objectives.
- (iii) To select the more probable alternative.
- (iv) To prefer the greater good (which helps to account for scheduling and adjusting ends to be mutually supporting.)
- (v) To order our objectives (by priorities) when they conflict.

Rawls’ LECTURES, *supra* note –, at 46.

<sup>124</sup> *Id.* at 436-37.

<sup>125</sup> See, e.g., Markus Dirk Dubber, *Recidivist Statutes as Arational Punishment*, 43 *BUFF. L. REV.* 689 (1995) (concluding that “three strike” statutes were passed and remain on the books despite the absence of a logical nexus between the punishment and its legitimate purposes).

<sup>126</sup> See note – *supra*.

assessments of the wrongdoer's conduct to naive, over-simplified truth claims about his character, which are reflected in our abbreviated semantics. FAEC-2 thus permits us to dismiss the wrongdoer's perspective and (worse, I think) to search for an audience with which we can share our self-certain truth about the wrongdoer.<sup>127</sup> The actor whose conduct we deplore – and whom we should hope to regenerate for the better – is now emotionally and actually forsaken.<sup>128</sup>

FAEC-3 uncovers the final act in the blaming process: the abandonment of an individual based on his bad character. This appears in part as the familiar practice of scapegoating, which succeeds emotionally (if not rationally<sup>129</sup>) when we target a wrongdoer's character based on the simplified, biased, and other-directed stories of wrongdoing we routinely tell ourselves.<sup>130</sup> The most troubling aspect of this destructive and unintended interpersonal process is that it provides us with license to deny our own responsibility in the wider web of relationships of which we and our targets are a part. The other becomes an *it* – *just the sort who* – and we are excused as a part of our social epistemology (what we take as a “usual, normal, and invisible” part of membership in our community<sup>131</sup>), and so we deny the “co-responsibility and contingency” in our lives and engage in a very troubling sort of moral accounting. When we blame without accounting as fully as we reasonably can for the “complex history of involvement in the events before, during, and after the ‘blameworthy’ deed,” we lose sight of the good and loving use of blaming.<sup>132</sup> This denial of biographical narrative (of the need to determine *who* this person is from a broader perspective) and the corresponding dismissal of co-responsibility becomes suspicious, I think, in direct proportion to the self-righteous certainty one brings to

---

<sup>127</sup> Williams, *supra* note # –, at 437. As Williams notes, the certainty of truth makes the actor's misdeeds all the more unthinkable to the observer who thus purports not merely to have an accurate picture of the reproved; he has a sufficient picture to condemn him with little if any appreciation for his perspective. “The speaker addresses the world, but refuses to engage with its realities: by what right, then, does he speak at all?” *Id.* at 438.

<sup>128</sup> For a dated but still helpful survey of early research on the variables that affect the biases in the social observer assessments see Gary Marks and Norman Miller, *Ten Years of Research on the False-Consensus Effect: An Empirical and Theoretical Review*, 102 PSYCHOL. BULL. 72 (1987).

<sup>129</sup> Understanding that “rationality” is itself a construct, here I'm opting for a definition that includes some measurable and pervasive capacity for empathy, which includes some sense of what Rawls describes as “reasonable principles . . . – those that regulate how a plurality of agents . . . are to conduct themselves in their relations with one another. Principles of fairness or justice that define the fair terms of cooperation are standard examples. So are principles associated with the moral virtues recognized by common sense: truthfulness, fidelity, and so on.” Rawls, LECTURES, *supra* note # – at 47.

<sup>130</sup> Williams, *supra* note # –, at 439.

<sup>131</sup> For a rich description of how this process operates, see Allen Buchanan, *Political Liberalism and Social Epistemology*, 32 PHIL. & PUB. AFF. 96 (2004).

<sup>132</sup> Williams, *supra* note # –, at 439.

process.<sup>133</sup>

Blaming as a social practice is deeply and too often tragically flawed; it fails to acknowledge an inalienable truth: that unaccountable contingencies mediate all inputs and outcomes in significant ways. As a result, compassion fails; we tend to assess the wrong committed based on the need to determine guilt in a simplified, biased moral accounting according to which every unwelcome outcome is equated with or projected onto an appropriately “guilty” individual. (This may account in part for the reluctance of many theorists to acknowledge the existence of moral luck.<sup>134</sup> It demands a willingness to invest substantial personal energy against deeply-ingrained retributive urges.) When we engage in the process of conjecturing the wrongdoer’s intentions in a moral audit to which we bring our own biases, we empower ourselves to disregard the fact that we bear some responsibility for the conditions that undermine for many the possibility of developing normatively appropriate practical reasoning skills.<sup>135</sup> When we disregard our own responsibility for the creation or continuation of the kinds of antecedent conditions and events that produce inappropriate conduct, we heap blame on the individual alone. But determinations of guilt are not simple matters of matching intentions with misconduct on a scale of comparative adversity,<sup>136</sup> something like Thompson’s connecting thesis, *CT*. It may be that this simplistic accounting of moral matching works at one level of attribution and of retribution – finding *a* causal agent and discharging the feeling of release among others<sup>137</sup> – but it fails to chronicle faithfully the myriad contingencies that “inhere between will and deed, between character and situation, between choice and result.”<sup>138</sup> We often opt for a form of reductive condemnation: We foist all such unwelcome contingencies on the individual, label him *bad*, and hope to be done with him.

In light FAEC, it is at least appropriate and fair to express some wonder about *who* the audience is for such unforgiving attributions of blame, *why* we choose this or that audience, and *what* we hope to accomplish on their behalf and ours when we engage in a

---

<sup>133</sup> *See id.*

<sup>134</sup> Among the many theorists who reject either the concept or importance of moral luck are Jeremy Horder, *Criminal Law: Between Determinism, Liberalism, and Criminal Justice*, 49 *CURRENT L. PROBS.* 159 (1996); Kadish, *Forward: The Criminal Law and the Luck of the Draw*, *supra* note # –; Moore, *Causation and the Excuses*, *supra* note # –, at 1092 (acknowledging the pressure of determinist forces but arguing that “moral responsibility for an action should be ascribed to an actor even when that action was caused by factors over which he has no control”); Morse, *The Moral Metaphysics of Causation*, *supra* note # –, at 883-86.

<sup>135</sup> *See* John L. Hill, (Note) *Freedom, Determinism and the Externalization of Responsibility in the Law: A Philosophical Analysis*, 76 *GEO. L. J.* 2045, 2059 (1988).

<sup>136</sup> Williams, *supra* note # –, at 439.

<sup>137</sup> *See* my *Some Thoughts on the Aesthetics of Retribution*, *supra* note # – *passim*.

<sup>138</sup> Williams, *supra* note # –, 441. For a description of the multiple levels of moral attribution in criminal law, *see* George Fletcher, *BASIC CONCEPTS OF CRIMINAL LAW* 82-4 (Oxford 1998).

process that so distances *us* (the good guys) and *them* (the bad ones) in a Whig culture formally committed to the idea of human improvement. It's as if we have forgotten that among the systemic issues this approach embodies is the failure to acknowledge a simple fact: If God had thrown the dice just one more time, the culprit being blamed *might* have been you or me.<sup>139</sup> Over such matters we have only good or bad fortune, and more or less of it. We certainly know that the world today is not always a just place, and we know that current notions of objectivity often are nothing more than yesterday's enlightened standards and tomorrow's *ancien regime*. We cannot claim widespread intersubjective normative agreement sufficient to inflict pain on all those wrongdoers who fail to meet our prescriptive standards if *we* neglect to reconsider the actual facts on the ground that inform normativity and, in the process, honor our obligation to situate all the players in a contingent world as best we can.<sup>140</sup> Blaming – and especially the process that transports our biases into legal practice – discounts the need for active participation by everyone in addressing the conditions that conduce to unwelcome conduct, and assumes a justness that simply does not exist in fact.<sup>141</sup> And facts are the lawyers' domain.

#### IV. Facing the Paradox of Evil

Accounting for the destructive features of blaming will advance our pursuit of a more accurate determination of desert; the process requires more care in combining situational and dispositional factors without bias; and that's a daunting task. As Amélie Rorty suggests, the task often insists that we go “all the way back.”<sup>142</sup> This retrieval process is not simply a matter of deciding *whether* an unsavory actor caused (even in a substantial way) an unwelcome event with sufficient cognitive capacity to attribute to him “knowledge” of its wrongfulness; or of finding the putatively applicable statutory *mens rea* mirrored in our reconstructed vision of the wrongdoer's brain at or attributably near *t*;<sup>143</sup> or, conversely, of

---

<sup>139</sup> Cf. Watson, *Responsibility*, *supra* note –, at 276-77 (ultimately rejecting the metaphor of reliance on a throw of the dice but understanding the sensation).

<sup>140</sup> Cf. Hume, *TREATISE*, *supra* note # –, at 489 (declaring that one “cannot [do] better” than to “consult” rational interests, where “better does not entail the best conceivable”). See generally Rawls, *LECTURES*, *supra* note # –, at 60. Rawls affirms that Hume's understanding of “conventions” (including those that move us toward the better) included striving for the “practically best scheme.” Writes Rawls: “[Hume] doesn't mean that it is the best scheme we can imagine, much less the best scheme allowing that human beings and our situation in nature might have been different. He means that it is the practically best scheme, accepting ourselves and our situation in nature as it is, without weeping or lament.”

<sup>141</sup> See Williams, *supra* note # –, at 442-43.

<sup>142</sup> See text accompanying note – *supra*.

<sup>143</sup> Fletcher, *BASIC CONCEPTS*, *supra* note # –, at 82.

determining whether some hyperpathology demands an excuse.<sup>144</sup> Rather, we also must commit to determining as best we can *how* he got to be that way; for without an answer to the *how* question, we have no hope of preventing the rush of what seems to be an unending supply of wrongdoers. This approach implicates a far richer view of both the potential and reality of human nature than settled jurisprudential doctrine recognizes, and it imposes attendant costs.<sup>145</sup> The practical difficulty with this suggestion is that it also produces an unsettling problem, accurately described as a paradox:

*The more egregious one's misconduct, the more likely he is to be damaged; the more likely he is to be damaged, the more likely he is to take advantage of an excuse; and the more likely he is to take advantage of an excuse, the more likely he is to avoid suffering the full penalties the law prescribes, which are themselves based on our collected folk wisdom.*

To the extent, therefore, that excuses are viewed as beneficial, because they permit the wrongdoer to escape some quantum of punishment, increasing their availability might produce perverse incentives to engage in the conduct that generates them.<sup>146</sup>

This result would surely present a serious problem if addouning for wrongdoing accurately wasn't consistent with our theological and moral commitment to compassion, and if the only alternative to the default recipe of imprisonment was permitting dangerous actors to roam the streets. But retributive imprisonment isn't the only alternative; civil commitment, or some more humane variant thereof, is an alternative to imprisonment.<sup>147</sup> Civil commitment is an alternative and it, too, depends on the good will of the entire system, from top to bottom.<sup>148</sup> But there's a deeper problem as well, and it concerns our

---

<sup>144</sup> Under the Model Penal Code, for example, an actor's cognitive process have completely decomposed, and he is (without differentiation) deemed insane, if he "believes he is squeezing lemons as he chokes his wife." M. P.C. cmt. to § 4.01, at 166.

<sup>145</sup> I have outlined the argument for symmetrical emotions in the context of Adam Smith's, *THE THEORY OF MORAL SENTIMENTS* (Prometheus Books 2000) (hereafter, "Smith, TMS") in *Some Thoughts on the Aesthetics of Retribution*, *supra* note # –, at 239-43.

<sup>146</sup> I thank the participants at a Washington University Law School faculty research program where I first presented some of these ideas for pointing out this potential problem.

<sup>147</sup> *See* note – *supra*.

<sup>148</sup> That this is an undoubtedly daunting task can neither be doubted nor used as an excuse to forego it. Foucault has described the layers of authority that must be brought on board to achieve any such revamping:

Small-scale legal systems and parallel judges have multiplied around the principal judgement: psychiatric or psychological experts, magistrates concerned with the implementation of sentences, educationalists, members of the prison service, all fragment the legal power to punish. . . . [A]s soon as the penalties and the security measures defined by the court are not absolutely determined, from the moment they may be modified along the way, . . . one is handing over to [these

razor-thin understanding of the propriety of blame; this understanding, too, ought to evoke some emotional tension.

Using the biographical highlights of the psychopath, Robert Alton Harris, as his test case, Gary Watson describes the source of a basic, intuition-based, moral conflict: “To be homicidally hateful and callous in Harris’s way [to eat the lunch of the innocent teens he has just brutally and senselessly killed] is to lack moral concern to an extent that comes as close to *entirely* as one can, and to so lack moral concern is to be incapacitated for moral community.”<sup>149</sup> This explanation, which is clearly consistent with one fork in the stream that our moral intuition follows, potentially exempts every hateful and callous criminal from moral disapprobation; and these exemptions might seem to arise more or less independently of any particular wrongdoer’s unique background. But in the case of Harris we know he had a hideous, loveless and brutalizing background.<sup>150</sup> In the case of Andrea Yates, we know she had suffered severe depressive conditions that erupted more than once into full blown psychotic episodes following childbirth. The conduct of Dahmer can be explained only as the manifestation of a catastrophic genetic defect: The taste for human flesh is well beyond any description of individual intentions one can fathom. Logically, he could not have caused his own desire to cannibalize his unsuspecting victims; only grave neurobiological imperfections and environmental catalysts can explain this desire.<sup>151</sup> Some form of incapacitation, therefore, is at least appropriate in some cases beyond any doubt

---

individuals] mechanisms of legal punishment to be used at their discretion.

Michael Foucault, *DISCIPLINE & PUNISH: THE BIRTH OF THE PRISON* 21 (Trans. by Alan Sheridan; New York: Vintage Books, 1978),

<sup>149</sup> Watson, *supra* note # –, at 274. What to do with psychopaths has evoked a variety of reactions from health care specialists, philosophers, and jurisprudence scholars. The premier researchers in this area of psychology are Hervey Cleckley, *THE MASK OF SANITY* (5th ed. 1988), and more recently Robert D. Hare. *See, e.g.*, Robert D. Hare and Michael J. Quinn, *Psychopathy and Autonomic Conditioning*, 77 *J. ABNORMAL PSYCHOL.* 223, 224 (1971), who suggests that psychopaths commit approximately 50% of all major crimes. R. D. Hare, *WITHOUT CONSCIENCE: THE DISTURBING WORLD OF THE PSYCHOPATHS AMONG US* 87 (Simon and Schuster, 1993). On Hare’s contribution to our understanding of psychopathy, *see, e.g.*, Ian Pitchford, *The Origins of Violence: Is Psychopathy an Adaptation?*, available on line at <http://human-nature.com/nibbs/01/psychopathy.html>. *See also* Elliott, *supra* note # –, at Ch. 5 (concluding that some of their misconduct is blameworthy despite acknowledging their lack of volitional control); Walter Glannon, *Psychopathy and Responsibility*, 14 *J. APPLIED PHILOS.* 263, 268 (1997) (noting that psychopaths lack the capacity to reason prudentially, which is a necessary condition for assessing one’s long term interests; Stephen J. Morse, *Rationality and Responsibility*, 74 *S. CAL. L. REV.* 251, 254 (2000) (stating that if “the emotional capacity for empathy” is deemed to be part of rationality, then psychopaths would not be considered rational; and concluding that psychopaths *are* “morally irrational” and thus should be excused).

<sup>150</sup> *See* note # – *supra*.

<sup>151</sup> The extent to which our genes are triggered and destroyed by the actual events in our lives is the subject of Ridley, *supra* note # –.

question. If necessary, so is indefinite isolation.<sup>152</sup>

The facts bearing on lack of capacity-in-fact and environmental opportunities represent a fundamental conundrum for the metaphysics that undergird our jurisprudence of punishment. For example, compatibilists, by definition, justify punishment within a determinist world.<sup>153</sup> For utilitarian compatibilists, what matters is that holding people responsible for their actions has efficacy in regulating behavior.<sup>154</sup> For others generally within this school, what matters is that blaming expresses and confirms community resentment.<sup>155</sup> In both cases the sought after penal goal is achieved whether or not the conditions of epistemologically-informed, full moral assessment are met. To the question, “Here we are today; what should we do from here?” different answers thus appears. For the regulator or expressivist, the answer is, “How he got to the point of blameworthy misconduct makes no difference.”<sup>156</sup> And for incompatibilists or hard determinists,<sup>157</sup> in contrast, Harris’s biography may prove their case: His background, if only in retrospect,<sup>158</sup> seemed to doom him; for if we view him as the mechanistic product of an egregious childhood, we cannot hold him responsible. (This does not mean that it is clear that Harris had to turn out as he did; his siblings did not become vile, hateful killers.) Given the level of abuse he suffered and what we know about the potential effects of abuse, and given Yates’ history of postpartum psychotic depressive episodes, our intelligible reactions to such cases

---

<sup>152</sup> I take no stance on the question whether such an individual *could* voluntarily opt for death. The topic is vast and, at present, beyond my ken. Certainly when incapacity is clearly and convincingly immeasurable by today’s standards, such decisions are even beyond what we now understand as minimal instrumental rationality. See note – *supra*.

<sup>153</sup> H. L. A. Hart, *Legal Responsibility and Excuses*, in PUNISHMENT AND RESPONSIBILITY 28, 44 (Oxford 1968). I have critiqued this view extensively in *Punishing Attempts*, *supra* note # –, at 971-77.

<sup>154</sup> See Richard W. Burgh, *Do the Guilty Deserve Punishment*, 79 J. PHILOS. 193, 194 (1982) (noting that there is a belief among utilitarians that “the greater good of punishment serves to mandate compliance with rules”).

<sup>155</sup> See, e.g., Joel Feinberg, THE MORAL LIMITS OF THE CRIMINAL LAW: HARMLESS WRONGDOING 149 (Oxford 1990) (stating that the relevance of moral blameworthiness “derives from its correspondence to an essential function of legal punishment which, as a symbolic device for expressing public reprobation, automatically stigmatizes the condemned offender”).

<sup>156</sup> See Watson, *supra* note # –, at 274.

<sup>157</sup> A useful discussion of the distinctions among compatibilists and determinists is found in P. S. Greenspan, *Behavior Control and Freedom of Action*, in MORAL RESPONSIBILITY 191, 191 n. 1 (ed. by John Martin Fischer; Cornell 1986).

<sup>158</sup> That is, from the perspective of theoretical reason, the lesson is obvious. Among several works in moral philosophy to examine the question of Kantian standpoint is Bok, *supra* note #– (distinguishing between theoretical reasoning, which seeks objectively accurate descriptions of the world, and practical reasoning, which asks what the agent should will on a given occasion).

ought to be and often is, “No wonder they turned out that way.”<sup>159</sup> There is no non-compassionate way to avoid these reactions.

Our ability to appreciate their emotionally and physically challenged backgrounds should cause us to react with appropriate moral uncertainty. We see their capacities for brutality and respond to them as the killers they are: They have inflicted unimaginable pain and suffering on wholly undeserving victims. But we also should see them as tragic figures caught within the web of uncontrollable antecedent events and conditions that trapped them within brutal lives, and so we sense that they too are victims. How many children suffer such a total absence of love and cruelty as did Harris and survives “normal”? Of Harris’s crimes, Watson writes, “Our sympathy toward the boy he was is at odds with the outrage toward the man he is.”<sup>160</sup> We still have good cause to fear him, therefore, and to fear his kind of violence, at least absent substantial successful treatment and, when necessary, a long period of confinement (and even indefinite incapacitation, provided adequate process).<sup>161</sup> But we should also be conflicted. Harris’s formative years were inflicted on him without any choice and yet he “unambivalently endorse[d] suffering, death, and destruction, and that is what (one form of) evil is.”<sup>162</sup> Following the birth of her firstborn, Yates suffered from a debilitating depressive episode through no fault of her own, and yet she viciously killed all her children. That we can understand these actions as predictable responses to horrifying conditions suggests both the fragility of our moral judgments and the need for compassion.

## V. On Coping With Our Irregular Sentiments

This work began by noting that the urge to blame is an aboriginal emotional sentiment. In his classic treatise, *THE THEORY OF MORAL SENTIMENTS*, Adam Smith observed that our judgments about blameworthiness (and praiseworthiness) hinge on three elements: (a) the *intent* with which an action is done, (b) the *action* that the intention produces, and (c) the good or bad *consequences* that follow from the action.<sup>163</sup> He argued that elements (b) and (c) – acts and consequences – were logically unrelated to our

---

<sup>159</sup> See Watson, *supra* note # –, at 275.

<sup>160</sup> *Id.*

<sup>161</sup> By this I mean that no one should be confined absent, at a minimum, clear and convincing evidence that the subject presents a genuine danger of substantial physical harm to himself or, more importantly, to others. The nature of the allegations, which must vary as a function of dangerousness and controllability (on balance), may impose a higher burden on the State.

<sup>162</sup> Watson, *supra* note # –, at 275.

<sup>163</sup> Adam Smith, *THE THEORY OF MORAL SENTIMENTS* 133 (Prometheus Books 2000) [hereafter, “Smith, TMS”].



assessments of blame.<sup>164</sup> Smith was presaging later “subjectivist” theorists who contend that harm should not be relevant to assessing blame:<sup>165</sup> the consequences, Smith opined, are “indifferent either to praise or blame.” He concludes that (a) – the “intention or affection of the heart” – is all that matters; and asserts strongly that this “maxim . . . [reflects] self evident justice.”<sup>166</sup>

Smith was a utilitarian.<sup>167</sup> He might have disparaged the retributive urge centuries ago as unnecessary idealism. Instead, he explained the strikingness of harmful consequences and our tendency to respond thereto with blame:

What chiefly outrages us against the man who injures or insults us is the little account which he seems to make of us, the unreasonable preference which he gives to himself above us, and the absurd self-love, by which he seems to imagine that other people may be sacrificed at any time to his convenience or his humor.<sup>168</sup>

Many of us experience something like this as an instance of “road rage,” the (sometimes nearly uncontrollable) anger that often ensues after someone cuts off or otherwise endangers another driver in traffic.<sup>169</sup> My normal (nearly-enraged but printable) reaction is likely to be: “How dare you *fr#%\*g* jackass imagine that I might be sacrificed at any time for your *fr#%\*g* convenience?” Smith’s analysis is right, of course. To deny this truth is to dismiss Hume wholesale.<sup>170</sup>

---

<sup>164</sup> *Id.* The irony that we require both (at least formally) as epistemological requisites to a procedurally sound conviction is beyond the scope of this article. See *Aesthetics of Retribution*, *supra* note # – (concluding that genuine harm to a first order interest will be a feature of our criminal jurisprudence for the foreseeable future).

<sup>165</sup> The “subjectivist” approach measures culpability based solely on the actor’s subjective intent. See, e.g., Joshua Dressler, UNDERSTANDING CRIMINAL LAW 378-80 (3d ed. 2001) (distinguishing between “subjectivist” and “objectivist” approaches to attempts).

<sup>166</sup> *Id.* at 134.

<sup>167</sup> In fact, Smith’s consequentialism in this instance operates indirectly; that is, he embraces a positive form of retribution, understanding that it would have a regulatory effect on deterrable conduct. See Russell, *supra* note # –, at 550-52. On the various modes of retribution, see J. L. Mackie, *Morality and the Emotions*, 1 CRIM. JUST. ETHICS 3, 4 (1982).

<sup>168</sup> *Id.* at 139.

<sup>169</sup> For a popular discussion of this phenomenon, see Patti Davis, *If You Can’t Say Something Nice . . .*, NEWSWEEK, Aug. 3, 2001, available at [www.keeppedia.com/ShowItemDetails.do?itemID=315977&extID=10043&data=road\\_rage](http://www.keeppedia.com/ShowItemDetails.do?itemID=315977&extID=10043&data=road_rage).

<sup>170</sup> Hume’s enduring achievement was, I think, to make us aware of the indefeasibility of our anthropological natures, and so he forces us to respect certain limits with which some Kantians hope to do away. Insofar as we (surely) *are* limited by our biological natures, we must debate the extent to which our understanding of “reason” and certain of the “calm” passions is premised on differences in our perceptual experiences and capacities,

But this tendency to transform harm committed into an over-ascription of blame that reflects the wrongdoer's evil character is not only unfair.<sup>171</sup> We should know that it is also counterproductive to the goal of an increasingly more sensible system of criminal justice because the extant view makes little room for meaningful efforts at rehabilitation. At least for me, it raises the never well answered old question that began this work: "How can we charge the reformatory view of punishment with disrespecting human personality, simply because it acknowledges humankind's animal inclinations (or our imperfect psychological natures) and hopes to use such information as a means to the good of his or her higher selves, remains incomprehensible. I say this understanding (as best I can) that no sane, fully flourishing person would advocate for a system that permits dangerous actors to roam the streets and threaten our families, our children and ourselves. My means and end are pragmatic *and* compassionate; the two cannot naturally (or factually) be oppositional. A humane practice of commitment, which avoids the regime of punishment, could provide an alternative,<sup>172</sup> because we need a regime that suspends blame, at least as a reflexive default social reaction.

The system I envision turns the normal process, which (if it addresses these queries at all) generally saves for the sentencing phase questions about (non-pathological) capacity-in-fact and opportunity so that certain decisions will be made on the front end of the process rather than the back-end. It retains a role for blame, but it does so only as part of a procedure that initially asks two questions and then operates on the two axes. The *X* axis measures capacity-in-fact: as best we can tell, does this individual have capacity for learning basic moral lessons? The *Y* axis examines past opportunities for moral education and the uses made in light of those opportunities.<sup>173</sup> What was possible where and when this person

---

together with our variably distributed verbal capacities and adaptive opportunities. Hume's vision of reason (as an instrument that not only cannot but *should not* master our passions) is set out in Hume, *TREATISE*, *supra* note # —, at 415 (contending famously that "[r]eason is, and ought only to be the slave of the passions, and can never pretend to any other office than to serve and obey them"). For a discussion of Hume's descriptive claim that reason is limited to (a) discerning patterns of events that imply a causal relationship and (b) figuring out the means to our desired ends, *see generally* Rawls, *LECTURES* 21-7 (examining, respectively, Hume's theory of the passions and the role of reason, and his critique of the rational intuitionists). On the obvious ambiguity in Hume's apparently normative declaration, *see* Annette Baier, *Master Passions* in *EXPLAINING EMOTIONS* 403 (ed. by Amélie O. Rorty; Cal 1980). It may be that Hume harbored a natural and healthy fear of unbridled, religiously-inspired views on the authority of rationality. *See, e.g.*, *TREATISE*, *supra* at 468-69 (dismissing the claims of the rational intuitionists that morality is discoverable by a certain kind of reason alone).

<sup>171</sup> It is also ironic that those theorists who attempt to explain our jurisprudence in terms of character overlook the bias that inheres in blaming even as they would do eliminate differences in penalties for attempts as opposed to completed harms. *See generally* my *A Jurisprudence for Punishing Attempts Asymmetrically*, *supra*.

<sup>172</sup> *See, e.g.*, Lawrence O. Gostin, *Public Health Theory and Practice in the Constitutional Design*, 11 *HEALTH MATRIX* 265, 308-10 (2001) (arguing for strict tests for civil commitment with meaningful treatment). *Cf.* *O'Connor v. Davidson*, 422 U.S. 563, 575-76 (1975) (holding that a state cannot confine a non-dangerous, mentally-ill person without providing meaningful treatment).

<sup>173</sup> I address this position in detail in *Blaming Angry Men*, *supra*.

matured, and what opportunities to advance were actually present? On the *X* axis, the state questions cognition and volitional control: we test to determine whether the wrongdoer *has* or does not (*lacks*) cognitive skills, and whether he *can* or *cannot* control his conduct within extant standards of control. If he fails on either dimension, he lacks adequate capacity-in-fact. On the *Y* axis, we examine as best we can, socioeconomic and educational opportunities. We ask, first, whether the actor *has* or *has not* have such opportunities, and if he, second, whether *has had* sufficient opportunity, whether he *did* or *did not* make some minimally sufficient use of them in the past. Under this proposal, meeting the current demands of minimal instrumental rationality<sup>174</sup> simply indicates that the actor should remain in the criminal justice system. For those unable even to effectuate a simple practical syllogism, cognitive competency remains an issue; that is to say, psychiatric treatment and incapacitation may still be required. And if has little or no volitional control (the psychopath being the model for such incapacity) he is relegated to a system that implements minimally coercive procedures to effectuate necessary constraint.

After this assessment is made, then a modified blaming and punishment practice takes over insofar as they remain practicable. Those who have ample cognitive capacity, affective control, and good fortune but fail to make use of them are, *prima facie*, the most culpable both morally and intuitively for they (when compared to those who are cognitively, emotionally, or experientially deprived) more likely *should have done otherwise*. (I also suspect their crimes are typically the more intelligible and deterrable as well.<sup>175</sup>) In contrast, those who lack both capacities and suffer bad antecedent luck are the least blameworthy for, *prima facie*, they likely *could not have done otherwise*. In the middle are those who possess sufficient cognitive skill and capacity-in-fact but lack the opportunity to adapt appropriately or lack the capacities but have been blessed with good opportunities; all may well require restraint, but little or no blame, and certainly not vengeful, retributive treatment. For all such groups, moral education during their period of incapacity is the only meaningful moral alternative if protecting future generations is our goal.

## V. Conclusion

I have tried to link our concepts of blame to the psychological phenomena that support it. I have examined the moral philosophy of blaming, which often produces the conclusion that the actor who causes harm is *bad*. In large part, that conclusion both inspires and rests on an abbreviated semantic description of character, which in turn both animates and supports the social and cultural biases we bring to the practice of attributing blame. We thus tend to overlook or devalue the striking nature of the conditions within which many wrongdoers reach maturity. This endemic tendency to over-attribute wrongdoing to the actor's bad character is at best unfortunate for it casts the wrongdoer as

---

<sup>174</sup> See note – *supra*.

<sup>175</sup> See, e.g., United States v. Bergman, 416 F. Supp. 496, 500 (S.D.N.Y. 1976) (noting that brief period of incarceration for wealthy, well-educated older man who plead guilty to multiple state and federal counts of criminal deliberate non-impulsive fraud “are among the most likely to be generally deterrable”) (citations omitted).

the target of our scorn, undermines efforts to make society safer by educating the actor, and removes us from responsibility for the social conditions that conduced to wrongdoing initially. At worst, it creates a subculture of Claggers.

In lieu of blaming as a routine practice applicable to all instances of wrongdoing, I suggest a new regime that suspends blaming, *prima facie*, pending an analysis of capacity-in-fact and opportunity. For those who lack capacity-in-fact and opportunity, the emphasis should be on education, although incapacitating the dangerous still may be warranted. For those who have ample capacity and opportunity, blame is appropriate for we have good reason to believe that they could have made choices other than those they made. In the end, the system I have outlined rests on compassion; the unwillingness to do to others what we would not want done to us, had God rolled the dice differently.