1. Introduction

In her stimulating essay, Professor Nita Farahany defines the new movement of “behavioral morality” and situates it within a taxonomy of movements that relate human biology to moral capacities and content. As Farahany’s taxonomy illustrates, there are numerous branches of inquiry that tread this field; yet, what may be genuinely novel in the work of behavioral moralists is their emphasis on relating the functions of and the activity within particular brain ensembles\(^1\) to specific types of moral cognition and action. While she is skeptical that behavioral moralists’ work on explicating causation will—or should—affect legal determinations of actors’ culpability, Farahany offers an avenue through which she believes behavioral morality could contribute to criminal law doctrine: she proposes that scholars working in behavioral morality should contribute at a more general level by refining the criminal law’s notion of the “reasonable person.” This is a doctrinally and scientifically practicable project and potentially a highly valuable one. Refinements of the reasonable-person standard could lead to more accurate judgments within the criminal law about the typical actor and could stretch beyond criminal law to other areas, like tort law.
While Farahany proposes the utterly reasonable project of rethinking reasonableness, this comment proposes a perhaps less prudent project of rethinking *unreasonable* reasonableness. This suggestion, although put lightly, is meant seriously. Behavioral moralists are interested in explicating the general neurobiology of morality, but many behavioral moralists are at least equally interested in explicating the morally salient differences between neurotypical and nonneurotypical actors. For behavioral moralists who work in this vein, rethinking reasonableness would be but a prelude to disaggregating “reasonableness” into several dimensions and then demonstrating how particular types of aberrant internal causation might compromise performance on those dimensions.

Scholars working in the behavioral moralist mode may be interested in showing how abnormal moral cognition leads to nonconforming social performance; they further may use this as a lens through which to explore the question whether abnormal biological causation should ever lead to legal and social exculpation. Put most generally, a central behavioral morality question for law would be “Can internal causes ever exculpate?” The predominant answer in both law and legal philosophy has been “no,” which is why Professor Farahany suggests that behavioral moralists direct their efforts to elucidating aggregate normalcy in the service of improving legal models thereof.

This invitation to rethink *unreasonable* reasonableness (or, really, the relationship between deviance and culpability) and Farahany’s invitation to behavioral moralists to rethink reasonableness both proceed from the same impulse: one that is exploratory and constructive toward the field of behavioral morality, while remaining agnostic about the general feasibility of or specific likely outcomes of these endeavors. This comment does not itself argue for how culpability determinations within the criminal law should be modified to account for biologically rooted moral deviance. Instead, it proposes criteria that behavioral moralist arguments in favor of differentiated legal culpability for nonneurotypical actors would have to satisfy, as well as identifies some of the practical, legal, and normative barriers to the development of such criteria. Specifically, I propose that behavioral moralists will have to identify differences in morally salient perception, affect, or cognition that are stable, atypical, severe, and relevant to the particular offense the
person committed. Furthermore, they must do so without merely reinscribing biological or social difference as disease ("otherizing"), since it is in theory possible to describe everything in physicalist language.

The project of rethinking unreasonableness relates to that of rethinking reasonableness in a further way: if typical moral cognition can be described biologically, as behavioral moralists propose, then certainly it should be possible to describe in biological terms deviant moral cognition (whether or not the biological level of description is the most relevant one, under the circumstances). Accepting a fully materialist account of cognition and behavior, is it possible to say that a person who is morally disordered is per se biologically disordered? More narrowly—and to take the causal arrow in the other direction—are there particular physical disorders that interfere significantly and specifically with aspects of moral cognition and performance?

These questions present as philosophical ones, but they are of particular interest to legal and policy scholars who focus on reducing antisocial behavior and promoting prosocial behavior. Interest in understanding the causes of lawbreaking conduct and the hope that causal understanding will lead to more effective legal policies and institutions have driven the law’s long and varied association with the biological and social sciences. While this new area of behavioral morality is arousing intense interest among legal scholars, what remains as vexing as ever is the appropriate relationship between causal explanation and legal consequences.

This relationship is particularly fraught in the area where applications for moral biology may be the greatest: criminal law. Questions of cause and consequence are most salient in criminal law because of the criminal law’s unique function of imposing not only penalties but also morally freighted judgment or “stigma,” a primary concern of Farahany’s piece. Imposing stigma seems just only when it is deserved. The criminal law—tracking common moral intuition—exonerates most harm-causing actions when the actor was not the volitional agent of the harm, even if he or she was its proximate cause. Other branches of law, from tort to contract to property, embody and express values of efficiency. Criminal law, too, has its utilitarian features, but it also deals heavily in the moral realm. Accordingly, in the criminal law, questions of the
relationship between cause and moral blame become questions both of fact and of value.

Part 2 expands the definition of behavioral morality offered in Farahany’s essay and then situates her specific area of concern within the broader behavioral morality field. In this section, I propose a tentative map or diagram of different scholarly positions relating internal causation with reduced moral culpability. Part 3 elucidates the important differences between behavioral moralists who take the position that any physical cause is exculpatory and those who argue that only aberrant physical causes are exculpatory. Turning in Part 4 to Farahany’s invitation to rethink reasonableness, I explore why the reasonable-person standard may be a productive and practicable target for behavioral moralists.

Part 5 suggests that behavioral moralists may be most interested in rethinking unreasonableness—in explaining and possibly exculpating deviant causation. This Part suggests a series of steps that may be necessary to identify legally relevant biological causes of deviance. The intuition of behavioral moralists that such factors should reduce culpability is normatively contentious, philosophically fraught, and, at this point, still scientifically tenuous. Yet, behavioral moralists working with new neuroscientific tools at least may be better equipped than prior generations of moral philosophers to make the argument that these kinds of causes are different from others and should matter morally and legally.

2. DEFINING BEHAVIORAL MORALITY

The potential impact of behavioral morality on law and related fields depends on what behavioral morality is. Farahany defines behavioral moralists as the set of thinkers who claim “that deviant behavior attributable to a physical cause is either less or not at all morally blameworthy.” She identifies two claims as central to behavioral morality, one positive and the other normative. The positive behavioral morality claim is that it eventually will be possible, on the basis of new brain science, to understand fully the causes of deviant and law-breaking behavior. The closely related normative claim is that knowledge of the physical causes of deviant behavior should lead to the replacement of existing punitive criminal and social responses to deviance with a therapeutic or medicalized
norm. On this view, causal explanations both will and should up-end determinations of criminal culpability (moral blame/censure) and criminal responsibility (whether one may be punished for wrongdoing). Farahany proposes that behavioral moralists are motivated to make such claims because they believe criminal offenders are unfairly stigmatized by the criminal law.

Farahany’s definition is provocative because it suggests that behavioral moralist thought is, at least in large part, teleological: that it is motivated by the goal of reducing stigma for offenders whose lawbreaking actions were physically caused. More radically, Farahany’s description points to a project among behavioral moralists to dismantle retributive elements of the criminal law by using physical causation to negate moral culpability.

Before exploring these aspects of behavioral morality and of Farahany’s argument, I want first to expand the definition of behavioral morality and to show where within the broader field the segment of behavioral morality on which Farahany focuses is located. I then propose that it would be useful to subdivide Farahany’s area of focus into two further areas: behavioral moralists who claim that all physical causal explanations are exonerating and those who claim that only aberrant physical causes are exonerating. This distinction is important because there are significantly different philosophical and legal stakes, as well as issues of proof, attached to each position. The distinction also matters to Farahany’s ultimate proposal, that behavioral moralists focus their attention not on the deviant actor but on refining the criminal law’s model of the reasonable person.

Behavioral Morality, Expanded

“Behavioral morality” generally describes the project of understanding moral reasoning as a product of and as being predicated upon human biology in interaction with its environment. Behavioral moralists mean this in more than the basic sense in which all materialists would agree that every thought, feeling, and behavior arises from and is effectuated through an organism’s biology. Rather, they explore the biological predicates of particular “moral” sentiments at the group and the individual level. Most behavioral moralists assert that forms of moral reasoning are intrinsic to the
biology of human beings; some would go further to claim that particular moral sentiments are preprogrammed into people via specific neurological structures and genetic encoding. Additionally, scholars in this area may investigate the impact of contextual factors and social cues on moral and normative judgments, as mediated through their effect on neurological activation patterns and neurochemical responses. Looking at the brain’s facilitation of moral or normative judgment in this more dynamic, “ecologically valid” fashion helps avoid (or at least moderates the risks of) acontextually reifying cognitive and affective responses that have social dimensions and culturally conditioned components.

William D. Casebeer describes behavioral morality as the attempt to “improve our understanding of the nature of moral theory and its place in moral judgment by treating morality as a natural phenomenon subject to constraints from and ultimately reduced to the cognitive and biological sciences.” Similarly, Patricia S. Churchland, one of the true innovators in biological philosophy, has described the field as focusing on “basic neuroscience research” questions, such as whether the brain contains anything like a “moral module”; whether, conversely, moral reasoning is broadly distributed throughout the brain; whether areas of the brain engaged by particular moral reasoning tasks are variable across subjects and cultures; and whether what philosophers have termed “moral reasoning” constitutes a biological natural kind or rather is—in terms of underlying biological processes, at least—a set of unlike phenomena. These scholars and their peers investigate the neural and physiological bases of moral cognition and the relationship between moral cognition and behavior at a level of inquiry simultaneously more basic and more abstract than that which might lead to direct legal conclusions about moral culpability for criminal wrongs.

Work on impaired subjects shows how damage to particular brain regions can alter morally relevant affect, cognition, and decision making. For example, there is now strong evidence that damage to brain regions “associated with empathy, rules compliance, and with moderating aggression or triggering inhibition” closely correlates with certain types of lawbreaking behavior. Such damage may result from physical trauma but may also be caused by the more chronic (and common) deformations of “childhood maltreatment, or other stressors such as post-traumatic stress of war.”
Researchers have identified that damage to areas of the orbital prefrontal cortex can result in “deficit[s] in acquisition of moral and social rules.” Thus, research is beginning to show which brain regions subserve a subject’s acquisition and comprehension of social and moral rules, which in turn influences perception, emotion, behavioral repertoire, and behavioral control.

Behavioral morality is but one of the areas of scholarship that focus on the relationship between human biology and morality; it is a close cousin of evolutionary morality. Evolutionary morality, as Farahany points out, looks at the general human capacity for moral judgment and posits that the “trait” of forming moral judgments must have conferred a survival advantage on early human groups. Evolutionary morality draws support from empirical work that shows constancy in certain moral judgments across populations and cultures and aims to describe the degrees of freedom (or constraint) that social institutions like the law may have on shaping moral intuition.10

Behavioral morality departs from evolutionary morality in ways that make it enticing to some legal practitioners and scholars. It can be more focused on present biological causes, rather than on prehistorical environmental constraints. Arguments that sound in behavioral morality may be more applicable or at least equally applicable to the individual subject as to the population group. Further, behavioral morality may (although it need not) focus on biological deviance, while evolutionary morality seeks to identify species-typical characteristics. For example, a behavioral morality argument might take the form of explaining at the level of mechanism how morally salient affect, cognition, and control become impaired in a person with a particular disorder, like fetal alcohol syndrome. The behavioral moralist argument might go on to draw conclusions as to whether aberrant biological causation should lead to a different ascription of moral blame and legal responsibility. Thus, behavioral morality lends itself more to direct legal advocacy than do related forms of scholarship.

In her definition of behavioral morality, Farahany emphasizes this advocacy dimension, postulating that behavioral moralists may be moved by emotional or social-justice concerns to exculpate people whose behavior can be traced to a physical cause. While scholars like Churchland and Casebeer define the field significantly
more broadly, Farahany’s definition captures the approach of scholars who focus on exonerating actors with identifiable biological causes of their antisocial or lawbreaking behavior.

**Behavioral Morality, Mapped**

As Farahany describes, certain behavioral moralists do assert that *all* physical causes of behavior are exculpatory (in the moral sense, if not the legal one). Others, however, claim that only *aberrant* physical causes, like structural or functional brain damage, are morally exculpatory. The premises and conclusions of these two groups differ in important ways.

It may be this latter group that is more motivated by the issue of stigma; the history of debate within criminal law about the punishment of people with biological deviance shows both that such people have been the targets of tremendous stigma and that advocates on their behalf are strongly moved by a perception that such stigma is unfair. Behavioral moralists may see a lawbreaking individual as the victim of disease, abuse, physical trauma, or one or more of a host of other biological or biosocial insults that impair brain function in particular ways; their intuition is that it is unjust to hold an actor responsible for deviance visited upon him by the confluence of bad environment and bad biology. In identifying stigma as a motivating concern of behavioral moralists, Farahany tacitly connects behavioral moralists to a deep and varied tradition of criminal law scholarship contending that the mentally impaired are overly punished and inappropriately stigmatized.

In identifying stigma as a motivating concern of behavioral moralists, Farahany tacitly connects behavioral moralists to a deep and varied tradition of criminal law scholarship contending that the mentally impaired are overly punished and inappropriately stigmatized.

In this essay, I propose a tentative map of scholars’ various positions on the relationship among moral culpability, legal responsibility, and physical cause (see figure 5.1). This map may help locate where different moral philosophy and behavioral moralist claims about the relationship between causation and moral responsibility lie. Mapping out the categories of arguments as to the relationship between cause and culpability will be helpful—I hope—in making more explicit the distinction Farahany points to among behavioral moralists. Some behavioral moralists argue that the fact of physical causation—all physical causation—is morally exculpatory. Others focus on abnormal causation only. These claims lie on different footings and lead to significantly different endpoints.
Figure 5.1

Legend

- A Deontological position
- B Behavioral morality (1)
- C Behavioral morality (2)

Can causation exonerate?

- Y
  - Types of causes matter
    - N
      - All causes are physical so “types of causes” is incomprehensible; OR all cause exonerates (B)
    - Y
      - Biological causation is special
        - N
          - Social or other causes equally or more important
        - Y
          - Biological aberration exonerates (C)

- N
  - Causation does not impinge on moral responsibility (A)

Exonerates morally

- Y
  - Exonerates morally and practically
  - N
    - Exonerates morally but not practically
    - N
      - Exonerates practically
3. Parsing Claims about Causation and Responsibility in Behavioral Morality

Revisiting Farahany’s definition of behavioral morality in relation to the map in figure 5.1, we find that she states that behavioral moralist thinkers claim “that deviant behavior attributable to [any] physical cause is either less . . . or not at all morally blameworthy.” This description of behavioral moralists would correspond to the [B] position in the map. She further notes that behavioral moralists might argue that “those individuals who fall outside the norm or range of average capacity to control their conduct ought to be held to a different level of social blame than those who can meet the capacities of the criminal law.” Behavioral moralists who make this claim would fall at the [C] position.

Although the [B] and [C] positions are significantly different, both are in opposition to the [A] position—the position that no internal physical causes are exculpatory. (The absence of minimal rationality may be exculpatory, but this test is causally agnostic; the presence or absence of minimal rationality can be assayed behaviorally.) The [A] position represents at a very general level the various schools of thought that assert that, although (or, agnostically, regardless of whether) all behavior is fully physically caused, legal and moral determinations of blame and punishment should be indifferent to causation. In the modern era, one could look to Michael Moore’s canonical tome Law and Psychiatry and Christopher Slobogin’s commentaries thereon as robust statements of this position. Most recently and vigorously, Stephen Morse has defended this territory, as has Michael Gazzaniga. These thinkers take the position, although on varied grounds, that causal arguments are not relevant to attributions of moral blame and legal responsibility.

Moving into the territory occupied by behavioral moralists, the [B] position represents scholars who argue that all behavior, including moral behavior, is fully physically caused and who conclude, on that basis, that all ascriptions of credit or blame are meaningless because a person does not have the opportunity to do otherwise. Joshua Greene and Jonathan Cohen, who are among the founders of experimental moral philosophy, represent this position. They assert that, because all behavior is fully caused, legal
accounts of an actor’s “choice” to “control” his or her behavior in any particular context are mere storytelling. This, they assert, is equally true of all actors: like some deontologists—although from very different grounds—they argue that distinguishing between actors’ relative moral culpability on the basis of whether or not they are neurotypical would not be a comprehensible project. However, their conclusion is the inverse of the deontologists’: no one, they contend, is the proper recipient of morally infused “blame” (or “credit”) for his or her actions. Rather, we need to eliminate folk psychological, retributive concepts and develop a purely harm-focused system. Such a system would evaluate the actor’s potential danger to society and confine, supervise, or otherwise treat him or her appropriately. This move would go to the very roots of the criminal system, both philosophically and historically.

The behavioral moralists at [C] differ from this position because of their emphasis on differentiating the moral and legal consequences of actions carried out by functioning moral agents from those of impaired agents. These scholars contend that moral behavior is fully enabled by a physical moral apparatus but that physical cause is not in itself inimical to moral blame. When functioning within a range of normalcy, an individual can be judged as a moral agent. Lacking neurological and associated behavioral function within a range of normalcy, however, he or she could be judged as if he or she were an intact moral agent. From the “is” of faulty causation, behavioral moralists of this kind move to the “ought” of reduced culpability. This group of behavioral moralists does not seek to eradicate the concept of blame from all criminal punishment; instead, it seeks to preserve and even extend the criminal law’s long-established distinctions in relative culpability based on excuses of mental abnormality. This may lead not to a conclusion of no moral agency but rather to a continuum: that agency and culpability determinations should be modified on the basis of or calibrated to the actor’s kind and degree of impairment.

While this continuum or calibration approach might sound cumbersome and impracticable, there is (some) legal precedent for such distinctions: competency to stand trial differs from competency to contract; both differ from competency to make a will. What constitutes sufficient competency for a given context involves determinations both of the different mental skills called for by the
task and of the different social-legal stakes, which might militate in favor of higher or lower competency standards than would be instituted by a purely physicalist system.

4. Farahany’s Invitation to Rethink Reasonableness

The heart of Farahany’s proposal is that behavioral moralists—both [B] and [C]—should focus on the average actor, not the deviant individual or class. She posits that, with “a more robust understanding of human behavior, the concept of reasonableness could be informed by both a scientific understanding of human behavior and the normative purposes of the criminal law.”21 Such a rethinking of reasonableness would be more practicable than attempting to mind-read as to individual actors; it also would contribute very broadly to the criminal law.

Farahany’s proposal that behavioral moralists rethink reasonableness has the dual benefits of being doctrinally constructive and scientifically practicable.22 At the level of doctrine, it is well accepted that the “reasonable person” is a convenient legal construct—not exactly a fiction, but not exactly not a fiction, either. As one of law’s major roles is to mediate in a pragmatic way between competing social goals, it is, of course, the prerogative of the law to create useful simplifications. Yet, the reasonable-person standard is one that frequently is ill drawn, not for reasons that serve a normative or pragmatic legal purpose but simply because it incorporates flawed assumptions. To the extent that the standard is inconsistent with real peoples’ behavior because it incorporates folk psychological mistakes and cultural baggage by flawed assumptions, rather than for good legal reasons (whether retributive, utilitarian, expressive, or legal realist), it is ripe for revision.

The reasonable-person standard was updated in significant ways in the late twentieth century to accommodate changed social norms and new research in psychology. For at least a hundred years, the strong belief existed in the culture and at law that the reasonable rape victim would struggle to the utmost, even to the extent of her life, and would report the attack promptly to legal authorities. Exhaustive research on the experiences and behaviors of and the social influences on real rape victims showed that victims tend not to engage in utmost physical resistance. Rather than
pursuing the prompt reporting previously deemed both expected and reasonable, rape victims frequently delay reporting or do not report the crime at all. Such research has filtered into law and public consciousness and has changed the general understanding of how the ordinary or reasonable person responds to sexual assault.

Revising reasonableness would have broad legal impact and would be culturally viable for legal scholars. Reasonableness standards are ubiquitous throughout criminal law and beyond. Modifying the law’s expectations of the reasonable person therefore would have broad impact. Further, as Professor Farahany adeptly explains, the task of revising criminal law mental states is effectively out of the hands of the scientists: the Model Penal Code (MPC) mental states of purpose, knowledge, recklessness, and negligence are purely behavioral and resist inquiry into the underlying causes of the individual’s acts. The MPC mental states take some of the color and romance out of criminal law: inquiries into killers’ “motives” and determinations of their “wickedness” or “moral depravity” live only in the realms of history and detective stories but are dead in the modern criminal law. Accordingly, modern criminal law is (largely) indifferent to whether a person killed because he was financially calculating, raised from birth to hate members of a particular group, or high on drugs. Farahany takes the prudent position that, because the MPC mental states are so fully entrenched in modern criminal theory and practice, talk of replacing them with richer models of cognition would be of only passing theoretical interest.

5. An Invitation to Rethink Unreasonableness

Professor Farahany has advanced the eminently reasonable proposal of rethinking reasonableness. In this Part, I offer a potentially unreasonable proposal to rethink unreasonableness. The latter is not offered as a substitute for the former, which is an important and promising endeavor. I offer it in the constructive and somewhat playful spirit of thinking through what it would mean to take seriously the differences between behavioral moralist [B] and [C] positions and what it then would take to construct a behavioral morality [C] position under which the criminal law might take account of deviant causation in evaluating culpability.
The different behavioral moralist positions [B] and [C] raise interesting questions for criminal law. Are some physical causes exonerating, while others are not? If so, is there an in-principle basis for the distinction or only ad hoc distinctions derived from norms and customs? Are there valid social and legal purposes for making distinctions as to causes, even if the biological and philosophical bases are lacking? Conversely, if biological and philosophical bases for such distinctions are found, might there still be social and legal grounds on which to refuse to make distinctions? Behavioral moralists [B] argue that all internal physical causes are exculpatory, so the challenge for scholars who argue [C] is to define precisely why, if we live in a world of complete physical causation, some biological causes but not others mitigate blame.

**Legal Deviance and Brain Difference**

How might one begin to construct the argument that certain kinds of physical difference—but not others—should result in different treatment under the criminal law? What, generally, would be a consistent and logically coherent way of constructing the [C] argument? Although I am agnostic about whether this would be a beneficial direction for the criminal law, I offer what follows in the exploratory and constructive spirit in which Professor Farahany offered proposals for a behavioral moralist rethinking of reasonableness.

As a first step, it would be necessary for behavioral moralists [C] to define what biological difference actually means in this context: there must be a generally acceptable and reasonably stable definition for the kinds of differences that scholars working in this behavioral moralist mode would argue to be exculpatory. Next, it would be necessary for such scholars to define empirically how particular kinds of biological difference relate causally to lawbreaking—an endeavor that is scientifically challenging, as well as politically contentious. After establishing these elements, it then would be necessary to move from the descriptive to the normative argument of why such differences should affect legal consequences.

In defining the relationship between biological difference and criminality, behavioral moralists [C] might now make a twofold contribution. First, a precise delineation of morally relevant brain
Amanda C. Pustilnik
aberration is needed. Second, neuroscience research on aberration and morality could develop a more precise description of how and why specific types of disorders increase the likelihood of an affected subject’s criminal justice involvement. What is needed is a compelling causal account of which neurological aberrations on average impair the subject’s morally salient affect, perception, cognition, or control—and why.

The first question for this research program likely would be: what kinds of neurobiological difference should—and should not—matter to moral culpability, and why? Not every newly discovered brain variation is a “brain disease”—despite media neurohype to the contrary.24 Every living human brain is individual, making the definition of “typical” and “aberrant” rise with scientific difficulty and potential normative judgment. Defining salient neurological difference for behavioral moralists’ purposes requires clear definitions of disease states and some method for defining disease that is relatively insulated from norms and biases.

The neurobiological aberrations that behavioral moralists [C] may count as exonerating should, I would posit, share at least the following characteristics: stability, atypicality, severity, and relevance. I will discuss each of these in turn.

Stability is a necessary characteristic of any type of exonerating brain aberration. Stability matters as it is one criterion that differentiates between disease conditions and fleeting situational responses. This is particularly important for behavioral moralists [C] in countering the claim that, since all behavior is biologically caused, no biological cause is meaningfully different from any other—and none should be exonerating. Behavioral moralists might distinguish between functional and situational responses that merely have negative social effects (e.g., a man gets angry at another for calling him names and punches him in the nose, an incident that could be fully described in biological terms but as to which the biological level of description is less powerful than the social-contextual description) and contrast them with persistent impairments that alter the actor’s morally salient perception, affect, cognition, or control.

Stability of a trait related to lawbreaking is on its own, however, insufficient. It must also be, among other things, atypical. The most obvious stable yet typical biological characteristic that cor-
relates highly with criminal justice involvement is sex: specifically, being male. Men commit almost all violent crimes; thus, maleness is the most significant “risk factor” for violent criminality. Yet, being male would not be exonerating under the principles sketched earlier because it is not aberrant; indeed, the very commonality of maleness shows that it is insufficient to relieve one of moral responsibility for violence, since moral rules have developed over time with the average man in mind.

Stable atypicality may begin to mark out the kinds of brain difference on which behavioral moralists [C] might focus, although stable atypicality is not independently sufficient; brain difference alone does not mean disease. People who are left-handed (like this author) certainly have brains that deviate from the norm. And, this set of brain differences produces reliably different behavior—unexpectedly throw someone a ball and the relationship between right-left brain difference and behavior will be obvious. But these differences are not plausibly related to impairments in moral reasoning or moral performance. The stable, atypical brain difference must be connected in ways that are describable at the level of mechanism to morally salient impairments in perception, affect, cognition, or control.

This last example of handedness shows an additional challenge in defining brain diseases (or disorders) that relate to morally relevant conduct. That problem is how to define intact or appropriate moral performance in ways that do not merely reflect contemporary social norms and inequalities. In other words, how would it be possible to screen out normative feedback and biases in the definition of “morally salient” brain disorder? For centuries, European and American culture deemed lefthandedness a sign of wickedness; use of the left signaled sympathy with the devil, the ultimate immorality. While behavioral moralists and evolutionary moralists might view morality as relatively constant across historical time periods and cultures, examples of significant moral change over time and place abound. Areas of significant moral dissensus include all types of sexual behavior—particularly, in contemporary times, whether homosexuality and bisexuality are immoral. Homosexuality may be the lefthandedness of today: it arises at least in part from brain difference, it is stable, it is not within the statistical norm, and it is equated by some with defective moral performance. Yet,
to equate these forms of difference with disease would be to engage in the worst kind of reverse reasoning: a type of behavior is defined as immoral and socially injurious, a brain-based difference related to the behavior is identified, and then the individual exhibiting the socially disfavored behaviors (and/or identity characteristics) is identified as morally diseased. Past efforts to ascribe biological bases to criminal lawbreaking provide abysmal examples of inscribing social difference as disease, so it would be particularly important to avoid such errors this time around.

The second step, elucidating the causal relationship between biological difference and lawbreaking, is equally important. This endeavor has engendered an extensive if uneven literature over the past century and a half. It has ignominious eugenic roots in the so-called biological criminology of the late nineteenth and early twentieth centuries, therapeutic roots in the asylum movements of the nineteenth century, and liberal roots in the mental health movements of the late twentieth century.

Behavioral moralists [C] may avoid the pitfalls of earlier attempts to biologize criminal deviance by reining in the field’s ambition: delimiting particular biological conditions that relate to criminal deviance, rather than attempting a complete biological description for all lawbreaking behavior, may be more productive and less strewn with landmines. Certainly not all—and probably not even most—people who exhibit antisocial and lawbreaking behavior have what could be described as mental illness or brain damage.25 Social norms, community influences, and demographic and situational factors may exert the greatest influences on whether a person commits a crime.26 While such factors all are moderated by and processed through an individual’s biological equipment, his or her response to such factors cannot be said to constitute a disorder that resides “within” the person’s neurobiology (or at least that does so in any kind of stable way). And, obviously, it is not possible to make the reverse inference that lawbreaking conduct (or other kinds of socially disfavored behavior) in itself indicates biological dysfunction.27

Recent work in law and neuroscience has posited a return—on totally different scientific footing—to notions of “biological criminology” based on brain difference, scrupulously scrubbed of the eugenic and racial dimensions of past work in this area. Research
on the relationship between mental health or mental difference and criminality shows that people who have certain forms of structural or biochemical brain dysfunction are significantly more likely to have particular limitations in perception, affect, cognition, and control that correspond with their particular form of brain damage. The incidences of mental illness and brain damage in the criminal system are multiples higher than in the general population; rates of mental illness among chronic recidivists are even higher.

What has eluded researchers until now is a causal explanation for how specific forms of brain difference influence or interfere with typical moral or social performance—with law-abidingness. And it is here—in the hard and detailed work of showing the mechanisms by which certain deficits impair an actor’s perception, affect, cognition, or control in ways that interfere with law-abidingness—that the major contribution could be made.

At this point, I anticipate the objection that connecting particular biological deficits with lawbreaking conduct would not only be difficult but pointless: law requires no more than minimal rationality of actors, and the causes of deviant behavior are legally irrelevant provided minimal rationality is present; presumed within minimal rationality is the capacity for ordinary self-control. This certainly is the position that [A] scholars would assert, as part of the claim that internal causes are not exculpatory. This view has a certain moral and philosophical purity, as well as practical appeal in that it leads to clear and implementable legal rules.

One possibility, though, is that work by behavioral moralists [C] and others may show that the concept of minimal rationality is not a standard of general legal accountability but rather a pragmatically useful legal fiction. I will attempt here an example of a domain-specific impairment that should be legally exculpatory in a particular instance. A person with a severe autism-spectrum disorder appears in court as a party or witness. This person has an average-high IQ and satisfies the requirements of minimal rationality: he has a basic understanding of cause and effect (he knows to get in out of the rain, for example), and has a cognitive understanding of reward, punishment, and consequences for action. When he addresses the court, the presiding judge tells him that he must look at the judge and must stop engaging in certain repetitive
physical behaviors. The actor not only does not stop the behaviors but indeed does them more, despite the threat of a contempt sanction. The judge eventually holds him in contempt of court.31

Autism’s deficits and differences are now fairly well defined (compared, at least, to those of other neurological developmental disorders); an autism expert could explain how the judge’s threat of contempt would have increased this actor’s anxiety, which would have had the effect of stoking—rather than suppressing—his repetitive physical behaviors. So, here we have an actor who has ordinary intelligence and minimal rationality; he understands what a “consequence” is, and he knows that he is supposed to do what a judge tells him. But the nature of his impairment makes his social performance less conforming, not more conforming, when he is threatened with the consequence by the judge. The judge in this example stands in for the proverbial “policeman at the elbow.” If the policeman at the elbow fails here, where else might it fail?

Resistance to Differentiated Culpability

If behavioral moralists [C] could accomplish the ambitious project outlined earlier, would the criminal law then embrace differentiated culpability based upon the morally salient and stable brain differences that behavioral moralists identified? The weight of past and current scholarly argument and popular sentiment suggests not. The strong resistance to differentiating culpability on the basis of neurological difference rests in part—but perhaps only in small part—on philosophical grounds. Those who take the [A] position, sketched out earlier, oppose such differentiation for in-principle reasons. Among the majority of commentators, legal scholars, lawmakers, and citizens, however, the resistance engendered by proposals for differentiated justice may be less philosophical than they are normative, emotional, and pragmatic.

Professor Farahany contends that behavioral moralists [C] are motivated by their perception of unfair stigma against people with neurological differences. Stigma certainly may play a strongly motivating role in legal discourses about the treatment of people with disabilities. But might not the opposition to the behavioral morality [C] position be at least as normatively driven? Major sources of resistance to differentiated culpability may include the role of
risk shifting in the criminal law, the immature and unreliable state of the science, pragmatic or efficiency-related concerns, and legal-doctrinal concerns. Finally, some of the resistance to behavioral morality [C] proposals may also be based in stigma against (or at least preconceptions about) lawbreakers, people with brain difference, or both.

Resistance to differentiated culpability may arise in part from the intuition that current criminal law regimes require a person with a behavioral difference to internalize the risks of his or her atypical behavior. Risk shifting is typically associated with tort law; it does not operate in any obvious way in the criminal law, framed as the criminal law is in the retributive language of moral blame and the utilitarian language of deterrence. And yet, criminal law implicitly addresses questions of which party should assume the risks and bear the costs of neurobiological difference. Let us take for example a person with severe fetal alcohol syndrome. This actor gets into a minor traffic accident; his neurobiological differences make it significantly more likely that he will become violent—in this case, he punches the other driver in the nose. We could think of the law’s current treatment of this actor, under which he receives no consideration for his mental difference, as functioning to internalize the risks of the actor’s disease to the actor. Why should the general right of innocent third parties to be unmolested in public be lessened by virtue of an impaired person’s instability?

The nascent state of brain sciences, too, poses a hurdle to the individualization of justice. Even in cases of gross anatomical abnormality, the link between brain difference and behavioral output in any one individual remains speculative—it is probabilistic and correlative, not individually provable. Some individuals experience major brain damage but are relatively behaviorally unimpaired. Others suffer minor head injuries and yet experience protracted mood and behavioral alteration, as with postconcussion syndrome; this may be particularly true in cases of low-level but diffuse brain damage that would not even be detectable with any current neuroimaging technology. Even in cases of demonstrable brain abnormality, the question of causation remains vexed: a defendant who presented evidence that his self-control was impaired due to postconcussion syndrome still would face the hurdle of linking the abnormalcy to his specific failure of self-control on the
occasion in question. This then raises the metaphysical question of the degree of control any individual ever has over his actions—a question that behavioral moralists may consider empirical but that other moral philosophers may consider not amenable to empirical determination.

Efficiency-related concerns may also underlie much resistance to individualizing justice on the basis of physical difference. These fall into two general categories, which relate to how to evaluate the offender and then how to dispose of him. First, most offenders will not have the benefits of an expert examination. Even if an expert does examine and identify in the defendant a significant structural or functional difference (and query what thresholds should be set for “significant”), it remains difficult and speculative to connect the difference to a legally relevant impairment.

If a causal connection is found, how should a court respond? The practical legal difficulties with individualized justice go further still, from evaluation to disposition: if richer mental state descriptions or other mechanisms for accounting for mental impairment were incorporated into the criminal law, what consequences would follow? The current repertoire of criminal sanctions is narrow: incarceration or supervised release. Justice individualized based on physical difference could make use of these alternatives, but this seems unlikely and unsatisfying. Expanded postconviction alternatives might be developed, like diversion to the mental health system. Yet, individualizing these postconviction alternatives would be as uncertain as individualizing the evaluative phase.

Considerations about the state of legal doctrine set up hurdles to individualizing justice on the basis of difference, as well. As Farahany aptly notes, modern criminal law does not inquire deeply into the mind of the offender; “motive” is merely a narrative device that ties together factual evidence. The evidence need only show whether the defendant committed the act and, if so, whether he knew what he was doing or disregarded a risk that his actions might cause a particular harm. These are the thinnest of mental states. Attempting to substitute these MPC mental states with richer mental-state descriptions—descriptions incorporating accounts of why the defendant acted as he did and whether his perception, cognition, and affect fall within a range of normalcy
—would require a sea change of almost unimaginable proportions across every area of modern criminal law.

Opposition to differentiated justice of the kind proposed by behavioral moralists [C] also engenders resistance on normative and emotional grounds. Empirical research demonstrates again and again that jurors and lawmakers believe brain difference should be an aggravating factor, not a mitigating one. The public greets claims of brain differences with deep suspicion, particularly since the insanity defense (which is now all but nonexistent) is mistakenly perceived as a get-out-of-jail-free card. Some of the hostility to claims of brain difference could be ameliorated through the development of more reliable neurological testing. Such a process of norm change through gradual assimilation of empirical findings is likely to be slow. This takes us back to Farahany’s important point of stigma: behavioral moralists [C] may (or may not) be overly influenced to “protect” people whom they perceive as victims of stigma in part because cultural stigma against people with mental illnesses is so significant.

Fairness arguments, too, have stood in the way of differentiated justice. A major argument against differentiating legal responsibility on the basis of mental disease or disability has been that doing so is unjustly underinclusive: childhood environment, trauma, abuse, demographic factors, and all of the other influences on individual development and on action-in-the-moment are just as physical and just as inexorable as mental disease. Contemporary work in cognitive neuroscience indeed is contributing to the body of evidence that growing up with severe environmental insults, whether chemical or situational, can trigger predispositions to mental illness or cause impairments equivalent to currently recognized mental illnesses. Thus, current legal excuses based on mental disease—like insanity—discriminate against people who are equally impaired but whose impairments arise from other sources. If it were to become accepted that biological impairment could, in principle, reduce moral culpability, then reduced culpability should apply to all people with valid physical conditions. Accordingly, any behavioral morality efforts to differentiate justice on the basis of physical cause may have to address the challenge that many factors traditionally labeled demographic, socioeconomic,
or emotional also could be construed as physical. Principled line-
drawing between these sources of difference may be the greatest
challenge of all.

6. CONCLUSIONS AND FURTHER DIRECTIONS

I share Professor Farahany’s conclusion that brain-based indi-
viduation of criminal justice is neither scientifically possible nor
normatively desirable. Differences in individual actors’ brains are
too hard to assess against any reliable baseline, and the criminal
justice system could not be so finely calibrated even if individual
brain readings were possible. Accordingly, Farahany appropriately
suggests that the right target of analysis for scholars in behavioral
morality is the typical community member, not the transgressor.

Picking up from Farahany’s work, this comment speculates as
to whether behavioral moralists could develop an intermediate po-

tion that focuses neither on the individual disordered offender
nor on the aggregate typical person but on aggregate kinds of ab-
normalcy—or “unreasonableness.” Instead of “cause” being an all-
purpose excuse—or no excuse at all—could particular aberrant
causes instead become the basis of taxonomies of exculpation, of
kind and degree of excuse? In addition to expectations we might
have of the typical person, based on work to date in behavioral
morality, what tailored expectations might we have of people who
meaningfully depart from the typical? I offer the suggestions in
this comment as initial suggestions of what it might mean for
scholars to think about whether deviant internal causation matters
in ways that are distinct from the general question of whether any
form of internal causation matters.

As we understand more fully how the lived environment shapes
people physically as moral agents and how biology lays the pre-
conditions for and establishes constraints upon moral cognition,
behavioral morality may require reconsideration of a society’s
moral responsibilities to its citizens and of people’s moral obliga-
tions to each other. Behavioral morality does not suggest that in-
dividual moral character is wholly inborn or that social mores are
biologically fixed. Rather, people have a fragile capacity for moral
flourishing that can be developed or deformed by social and physi-
cal environment.
One important criticism of various historical and contemporary efforts to ground criminal law on a neuroscientific footing—or to develop a modern “biological criminality”—has been that they focus too narrowly on the individual in isolation—as if individuals and their constituent bits of brain were free-standing, stable entities. What we learn from behavioral morality, developmental neuroscience, behavioral economics, and situationist psychology is this: individuals’ biological capacities and moral intuitions both arise from physical primes and also are shaped, constrained, and given content by the physical and social matrices in which individuals are embedded. Some biological conditions are so extreme that they render the individual indifferent to situational and social factors; a child with severe brain damage will not flourish beyond a certain level regardless of the degree of environmental enrichment. But the converse is not true: a person who starts out perfectly intact is vulnerable to becoming grotesquely cognitively, affectively, and morally deformed under the wrong environmental and social conditions. In other words, we may have more downward than upward variability—or vulnerability. The moral intuition that such factors should reduce culpability is a contentious one, by no means universally shared. But behavioral moralists at least may be equipped to make the argument that these kinds of causes are different from others, and should matter morally and legally.

NOTES

Thanks to Richard C. Boldt, Danielle A. Citron, Nita A. Farahany, James E. Fleming, Oliver Goodenough, and Sandy Levinson. Jonathan Tippens provided invaluable research assistance.

1. By “ensembles,” I mean regions of the brain that become active in concert.

2. E.g., as when the actor is impelled by an external force or suffers a nonnegligent accident.

3. This is not to say that criminal law is the exclusive branch of law that expresses social norms. Law inevitably reflects the norms of its drafters or constituents; the apotheosis of economic values in various branches of law may itself be seen as normative—and different schools of economic thought may have different normative valences. However, I am
distinguishing here the criminal law’s peculiar function of passing moral judgment. The purpose of a judgment in contract or tort is to reallocate resources, not to brand the loser a “bad man”; the purpose of a judgment in criminal law is, in part, to hold up the actor to public disapprobation.


14. There are numerous, robust philosophical arguments as to why causal stories should not impinge upon moral blameworthiness; it is the purpose of this comment not to survey or extend them but merely to note their place in the debate about the relationship between causal arguments and the claims of behavioral moralists.
18. Gazzaniga, The Ethical Brain.
19. They further contend pragmatically that folk psychological concepts of “choice” and “volition,” whether or not they are true, are indispensable for ordering individual and social behavior. This makes sense intuitively; additionally, a growing body of empirical literature supports the conclusion that people behave better when they believe that they have control over their behavior. See, e.g., Kathleen D. Vohs and Jonathan W. Schooler, “The Value of Believing in Free Will: Encouraging a Belief in Determinism Increases Cheating,” Psychological Science 19 (2008): 49 (finding that subjects primed with information about free will were less likely to cheat on an experimental task and those primed with information about determinism were more likely to cheat); Roy F. Baumeister et al., “Pro-social Benefits of Feeling Free: Disbelief in Free Will Increases Aggression and Reduces Helpfulness,” Personality and Social Psychology Bulletin 35 (2009): 260.
22. The scientific aspects of this proposal are fully addressed in Farahany, “Law and Behavioral Morality.”
23. Morse, “Brain Overclaim Syndrome and Criminal Responsibility.”
24. Two of my favorite recent examples: music is potentially addictive because the experience of listening to it releases the same brain chemicals
as drugs and sex! Yes: the dopaminergic system is involved in pleasure and reward, but that does not mean that music = drugs or that listening to music is the brain disorder of addiction (although the triumvirate of sex, drugs, and rock and roll must have been catnip to science reporters, whose beat is usually much drier). Further, rudeness is a “neurotoxin,” and people who are rude have been “infected.”

25. Some scholars who work in the behavioral moralist vein, broadly construed, have asserted that brain dysfunction can be inferred from the mere fact of social deviance; there is little evidence for drawing the causal arrow in that direction and many issues with such claims. For a critique of that position, see Pustilnik, “Violence on the Brain.”

26. See, e.g., Jon Hanson and Michael McCann, “Situationist Torts,” Loyola of Los Angeles Law Review 41 (2008): 1350; Nita A. Farahany and James E. Coleman Jr., “To Know the Criminal from the Crime,” Law and Contemporary Problems 69 (2006): 128. These social factors may also be the subject of biological inquiry; situational and environmental factors affect behavior only insofar as the individual takes them in and processes them affectively and cognitively against the background of memory and predisposition.

27. For a debunking of this reverse inference argument, see Pustilnik, “Violence on the Brain.”

28. Ibid.

29. Ibid., 207 (citing Nathaniel J. Pallone and James J. Hennessy, “Brain Dysfunction and Criminal Violence,” Society 35, no. 21 (1998): 21; Richard E. Redding, “Why It Is Essential to Teach about Mental Health Issues in Criminal Law (and a Primer on How to Do It),” Washington University Journal of Law and Policy 14 (2004): 407, 408–10). This is not to say that most people with severe mental illnesses in fact cross paths with the criminal justice system; they do not. And, there are many factors that influence why some people with severe mental illness do come to the attention of the criminal system—not least that the criminal system has become a hospital system of last resort for the uninsured and underinsured with mental illnesses. So the prevalence of people with mental illness and brain damage in the criminal system is not due to biological causation alone; it rests as well on interrelated political and economic factors. Nevertheless, the economic and political factors that contribute to concentrating people with mental illnesses or brain damage in the criminal system themselves show the importance of biology: in the absence of the biological problem, people with these conditions would not need political and financial resources to keep them out of prison.

30. Again, as discussed earlier, those in the [A] position may assert that only minimally rational agents may be held culpable. The focus of that
inquiry is the presence or absence of minimal rationality, not the types of rationality deficits or the causes thereof. Such an inquiry as to minimal rationality, then, is cause-agnostic. Moral behavior is fully enabled by a physical moral apparatus, which, when functioning within a range of normalcy, permits the individual to be judged as a moral agent. Farahany proposes that a person should not be judged morally if he or she is so impaired that “attributions of human agency” would be inappropriate. Farahany, “Law and Behavioral Morality,” 151. I interpret this standard to be equivalent to the standard of minimal rationality.

31. We could say that this is an unrealistic example because a judge would never be such a jerk; however, it would be within the scope of the judge’s power to act in this way, and only discretion would prevent it.

32. Occasionally, the mental state of criminal negligence suffices for conviction. Criminal negligence is the culpable absence of a normatively mandated mental state of due care.