

Winter 2019

Motus Animi in Mente Insana: An Emotion-Oriented Paradigm of Legal Insanity Informed by the Neuroscience of Moral Judgments and Decision-Making

Federica Coppola

Follow this and additional works at: <https://scholarlycommons.law.northwestern.edu/jclc>

 Part of the [Criminal Law Commons](#)

Recommended Citation

Federica Coppola, *Motus Animi in Mente Insana: An Emotion-Oriented Paradigm of Legal Insanity Informed by the Neuroscience of Moral Judgments and Decision-Making*, 109 J. CRIM. L. & CRIMINOLOGY 1 (2019).
<https://scholarlycommons.law.northwestern.edu/jclc/vol109/iss1/1>

This Article is brought to you for free and open access by Northwestern Pritzker School of Law Scholarly Commons. It has been accepted for inclusion in Journal of Criminal Law and Criminology by an authorized editor of Northwestern Pritzker School of Law Scholarly Commons.

CRIMINAL LAW

MOTUS ANIMI IN MENTE INSANA: AN EMOTION-ORIENTED PARADIGM OF LEGAL INSANITY INFORMED BY THE NEUROSCIENCE OF MORAL JUDGMENTS AND DECISION-MAKING

FEDERICA COPPOLA*

Legal insanity is deeply rooted in an intellectualistic conception of the capacity for moral rationality. The vast majority of insanity standards essentially consider the integrity of the defendant's cognitive faculties at the time of the offense. However, the cognitivist model of legal insanity collides with the body of neuroscientific and behavioral literature about the critical role of emotions in moral judgments and decision-making processes. Drawing upon this scientific knowledge, this Article reforms the intellectualistic substance of the capacity for moral rationality that underlies the insanity doctrine by including emotions in its relevant psychological set. Hence, it provides a revised model of legal insanity, one that gives more prominence to individuals' emotional faculties in relation to the crime committed. The analysis highlights that the legal reconsideration of the role of emotions within the capacity for moral rationality turns the insanity defense into a tripartite, more dimensional test—one inclusive of emotional, cognitive, and volitional prongs. Normative arguments in support of the proposed alternative paradigm of legal insanity are illustrated and discussed.

* Postdoctoral Research Scholar, Presidential Scholars in Society and Neuroscience (PSSN), The Center for Science and Society, Columbia University in the City of New York; Lecturer in Criminal Law and Neuroscience, Columbia Law School; University of Bologna Law School, J.D., 2010; European University Institute, LL.M. in International, European and Comparative Laws, 2014; European University Institute, Ph.D in Law, 2017.

TABLE OF CONTENTS

INTRODUCTION.....	2
I. THE RATIONALISM BEHIND THE INSANITY DOCTRINE.....	8
A. Cognition and the Cognitive Prong of Insanity Tests	13
B. Cognition and the Volitional Prong of Insanity Tests	18
II. THE NEGLIGIBLE ROLE OF EMOTIONS	22
A. “Moral Insanity” and the Lack of an Emotional Capacity Test	24
B. Lack of Self-Control and Emotional Disturbance	28
III. EMOTIONS, MORALITY, AND ANTISOCIAL BEHAVIOR: INSIGHTS FROM NEUROSCIENCE.....	30
A. “Knowing” Without “Feeling” Has a Negative Impact on Moral Judgment and Decision-Making	31
B. Self-Control Abilities Also Depend on Emotions	43
IV. A TRIPARTITE TEST FOR LEGAL INSANITY	49
A. Limiting the Scope of the Cognitive Prong.....	50
B. Including an Emotional Prong.....	51
C. Integrating Emotions in the Substance of the Volitional Prong.....	55
D. Rethinking Diminished Capacity as Generic Partial Insanity	57
V. NORMATIVE ARGUMENTS SUPPORTING AN EMOTION-ORIENTED MODEL OF LEGAL INSANITY	60
A. Personal Guilt.....	62
B. Culpability-Based Retribution.....	63
C. Rehabilitation	65
CONCLUSION	69

INTRODUCTION

The relationship between neuroscientific disciplines and legal insanity has never been simple. Many books and articles have been written, many conferences have been held, and many contrasting views have been proposed, but the debate continues. On the one side of the spectrum, some authors have called for neurological defense on the grounds that brain diseases may excuse the crime.¹ At the other end of the spectrum, some

¹ See, e.g., Richard Redding, *The Brain-Disordered Defendant: Neuroscience and Legal Insanity in the Twenty-First Century*, 56 AM. U. L. REV. 51, 53(2006) (arguing in favor of a return to control tests for insanity “that comport with modern neuroscience research on the role of brain dysfunction in impulsive criminal behavior”).

scholars have expressed concern that brain images and scans can actually mislead juries, giving them the wrong impression that the brain is wholly responsible for human behavior, which may allow for criminal conduct to be excused based on any brain abnormality in the defendant.² The lack of resolution of these disputes makes ambiguous the contribution of neuroscience to legal insanity, even at the theoretical level. This Article attempts to fill this gap and proposes one possible approach by which neuroscientific knowledge may plausibly contribute to a rethinking of the insanity doctrine without causing any dramatic upheaval to the nature of culpability and criminal responsibility.

Brain mechanisms do not alone account for an individual's (lack of) culpability. Culpability and insanity are not neuroscientific concepts, nor can they be localized in certain neural patterns.³ Yet, although brain mechanisms cannot provide an answer to normative questions about culpability and criminal responsibility,⁴ these physical features may become integral to discussions of culpability (and the lack thereof), as long as they contribute to a better understanding of the processes that underlie the capacities necessary for one to be considered culpable. This consideration of brain mechanisms, once again, does not equate to attributing a normative significance to neuroscience, nor to claiming that neuroscience could erode the nature of culpability and criminal responsibility. Rather, neuroscientific information can be used as a source of knowledge to improve the accuracy of the legal-psychological assumptions that support notions of culpability.

² See, e.g., Stephen J. Morse, *New Neuroscience, Old Problems*, in *NEUROSCIENCE AND THE LAW: BRAIN, MIND, AND THE SCALES OF JUSTICE* 157 (Brent Garland ed., 2004); Michael Pardo & Dennis Patterson, *Philosophical Foundations of Law and Neuroscience*, 2010 U. ILL. L. REV. 1211, 1213 (“At its most general level, law regulates human behavior. Human action *tout court* is just one of many issues within the scope of the neurolaw literature. Moreover, legal judgments are made by human decision makers, who act and decide based on reasons, which is just more behavior to be reduced within a neurolaw framework. Given the strong claims made on behalf of neuroscience, coupled with the growing enthusiasm for the enterprise, careful scrutiny is warranted.”).

³ See, e.g., Stephen J. Morse, *Brain Overclaim Syndrome and Criminal Responsibility*, 3 OHIO ST. J. CRIM. L. 397, 405 (2006) (“The criteria for responsibility are behavioral and normative, not empirically demonstrable states of the brain Brains are not held responsible. Acting people are. To believe that brain evidence has more than simple evidentiary value for assessing responsibility is to misconceive the criteria for responsibility.”).

⁴ See, e.g., Dean Mobbs et al., *Law, Responsibility, and the Brain*, 5 PLOS BIOL. 0693, 0696–97 (2007) (observing that “[d]etermining criminal responsibility is a normative legal conclusion, not an empirical factual one, made in the context of a variety of often conflicting aspirations. Therefore, even the best neuroscientific study can only afford factual evidence to be weighed alongside . . . normative considerations, rather than actually resolve the legal question as to which the factual evidence is relevant.”).

Over the past thirty years, neuroscience research has greatly advanced our understanding of the dynamics that underlie decision-making processes leading to moral conduct. One of most relevant insights emerging from this research concerns the critical role that emotions and emotional processes play either in informing or in hindering moral decision-making.⁵ In confirmation of this insight, brain-imaging studies on specific psychiatric populations characterized by marked antisocial tendencies have found links between these conditions and abnormal structure or functioning of the same socio-emotional brain circuits that appear to be significantly involved in moral decision-making.⁶ Altogether, consistent with behavioral studies, research in neuroscience supports the view that emotions are crucial mediators for moral behavior; that is, moral behavior also depends largely on proper and balanced emotional functioning.⁷

This Article specifically uses this branch of neuroscientific knowledge to revise the cognitivist model of the capacity for moral rationality, which lies at the core of the insanity defense. It provides an alternative model of legal insanity: one that gives more prominence to individuals' emotional faculties in relation to the crime committed. Additionally, it offers several arguments for why an emotion-oriented model of legal insanity—informed by scientific knowledge—is normatively plausible.

The argument offered here proceeds as follows. Part I traces the cognitivist model of insanity in contemporary criminal law. It begins with a preliminary discussion of the intellectualistic view of the capacity for moral rationality, which forms the benchmark of culpability and criminal responsibility. As will be made plain, the capacity for moral rationality consists of one's ability to engage in instrumental practical reasoning dictated by moral reasons. Importantly, in the eyes of the law, the capacity for moral rationality is entirely governed by cognitive faculties. Therefore, agents may be considered culpable as long as they possess intact cognitive faculties that enable them to know or understand the meaning of their unlawful conduct and willfully choose to engage in that unlawful conduct accordingly. In sum, cognition is the only mental dimension that defines the legally relevant mind. Part I subsequently explores how the intellect-based understanding of the capacity for moral rationality is reflected in the insanity doctrine. By analyzing formulations of insanity standards, it highlights that insanity tests are fundamentally focused on the evaluation of a defendant's cognitive faculties at the time of the offence. On the one

⁵ See *infra* Part III.

⁶ *Id.*

⁷ *Id.*

hand, cognitive defects are considered responsible for a defendant's lack of knowledge, or understanding, of the factual and moral meaning of the offense (cognitive prong of legal insanity). On the other hand, by impairing defendants' capacity for comprehension, cognitive defects are also assumed to affect their capacity to control their impulses (volitional prong of legal insanity).

Part II provides further support for the criminal law's adherence to an intellectualistic model of legal insanity by analyzing the negligible role that emotions are afforded within the evaluation of insanity. In line with this rationalist perspective, criminal law manifests a view that emotions make no positive contribution to moral rational reasoning. It presupposes that emotions are mental occurrences that, when excessively intense, can provoke sudden loss of control.

The negative relationship between emotions and the capacity for moral rationality is echoed in insanity standards in two ways. First, insanity standards do not provide an emotional capacity test, which is to say, a test measuring a defendant's capacity to emotionally appreciate the moral significance of the offence. Second, insanity standards give prominence to self-control impairments as long as they are linked to a defect of cognitive faculties. On the contrary, self-control impairments arising from emotional disturbance are usually considered as mitigating circumstances, to be considered in the sentencing phase or as limited diminished-capacity conditions, such as the common law "heat of passion" and the Model Penal Code's (MPC's) "extreme emotional disturbance" (EED). The diminished weight of emotion in such an evaluation, as I contend, is for one fundamental reason: because emotions are not treated as mental factors that contribute to one's capacity for moral rationality, a lack of self-control due to emotional impairment is not viewed as the kind of moral rationality defect that can justify an excuse.

Part III measures the rationalist model of legal insanity against neuroscientific insights into the role of emotions in moral judgments and decision-making. Its objective is not to carry out a detailed literature review of neuroscientific studies. Rather, and more narrowly, it aims to use relevant neuroscientific literature to emphasize two main mistaken assumptions about moral decision-making and behavior emerging from the current intellect-based model of legal insanity. First, it outlines that in moral judgments and decision-making processes, emotional faculties play a role equally critical to that of cognitive faculties. Notably, it emphasizes that emotions and emotional faculties influence moral judgments and

decision-making at both subconscious and conscious levels⁸ and, thus, that cognitive faculties alone cannot give rise to moral decisions without

⁸ In common language, emotions and feelings are used interchangeably. An emotion is usually referred to as something that is felt. People generally call an emotion their feeling of fear, happiness, anger, and so on. However, it is important to note the distinction that neuroscientists draw between emotions and feelings. Emotions can be defined as “the process[es] by which the brain determines or computes the value of a stimulus.” JOSEPH LEDOUX, *SYNAPTIC SELF: HOW OUR BRAINS BECOME WHO WE ARE* 206 (2002). In contrast, feelings are the subjective experience, or awareness, of said emotional responses. While emotions are mostly unconscious, feelings imply some degree of awareness (i.e., consciousness). When a stimulus occurs, people react subconsciously with their emotions. Only after the emotion—that is, the automatic response to that stimulus—has occurred do people become aware of it. This state of awareness is what transforms an emotion into a feeling. In other words, it is only when people get to a stage of awareness of the processes activated by an emotion that they have a feeling. Although there is a common view on what kind of states emotions and feelings are, as well as on the fact that both emotions and feelings do influence decision-making and behavior, neuroscientists hold heterogeneous positions on the exact relationship between emotions and feelings. See Joseph LeDoux, *Feelings: What are They and How Does the Brain Make Them?*, 144 *DAEDALUS, J. AM. ACAD. ARTS & SCI.* 96 (2015) (suggesting that emotions and feelings serve radically different functions, and thus cannot be placed on the same level. Emotions are adaptive responses to critical environmental challenges. An emotional reaction is a pivotal behavior of all organisms, both human and nonhuman. Emotions form automatic behavioral responses, motivational states to external stimuli that serve survival functions. However, emotions do not contribute to the emotional life of an individual. On the other hand, feelings are the aware and self-reported experience of an emotional response. Feelings are a matter of consciousness. Consciousness, and therefore prototypical cognitive systems and functions, are what make us emotional); cf. ANTONIO DAMASIO, *THE FEELING OF WHAT HAPPENS: BODY AND EMOTIONS IN THE MAKING OF CONSCIOUSNESS* 36 (1999) (arguing that feelings are natural evolutions of emotions. Emotions affect the mind when they evolve into feelings. Because feelings are the natural sequence of emotions, one’s feeling of a given emotion is ultimately the emotion per se. Thus, because each feeling is the natural corollary of a respective emotion, *the two things can be referred to by using same names*. For instance, *fear* can be both an emotion—i.e., a body state of change triggered by an external threat—and a feeling, i.e., the conscious perception of the body change. While emotions are evolutionary adaptations, unaware and embodied states that trigger physiological responses to external stimuli, feelings are nothing more nor less than the conscious perceptions, the lasting memory, a neural and mental representation of emotions. According to Damasio, while “emotions-proper” are always unconscious, feelings may or may not involve consciousness. He spots three stages of emotional processing along a continuum: “a state of emotion, which can be triggered and executed nonconsciously; a state of feeling, which can be represented nonconsciously; and a state of feeling made conscious, i.e., known to the organism having both emotion and feeling”). To put it simply, emotions provoke changes in the body. These bodily changes are projected and mapped in the brain. Bodily changes may remain non-conscious or may be experienced consciously as ‘feelings.’ Therefore, emotionally-salient stimuli may lead to feelings indirectly by triggering an emotion that causes a change in body state which is subsequently ‘felt.’ The line between the emotion and the feeling may thus be very blurred. As feelings are the natural sequence of emotions, we may say that our feeling of a given emotion is ultimately the emotion per se. See also Ralph

emotional influence. Second, the Article suggests that self-control abilities depend on their own mechanisms, encompassing many distinct (and dissociable) cognitive and socio-emotional processes. As such, people's capacity for self-control does not necessarily depend on the cognitive ability to know that a certain action is wrong.

In light of the highlighted scientific insights, Part IV draws up and proposes a tripartite model for legal insanity which accounts for the relevance of emotional factors. While the cognitive prong of the insanity tests remains essentially unaltered, the Article first advocates for the inclusion of an emotional capacity test—a test measuring defendants' capacity to emotionally appreciate the moral significance of their conduct—in insanity standards. Second, it advocates for the recognition of an independent volitional prong for the test, to measure the defendants' abilities to make decisions and exercise self-control, regardless of their intellectual ability to tell right from wrong. Furthermore, Part IV analyzes the consequences that the expansion of the substance of the volitional prong to also incorporate emotional components has for the diminished-capacity doctrine, as it is regulated by both the common law "heat of passion" and the MPC's EED standards.

Part V illustrates several arguments that support the normative plausibility of the new model developed for the insanity defense. In particular, it argues that an emotion-oriented model of legal insanity not only better complies with the principle of personal guilt, but also meets the aims of two major justifications for punishment, namely culpability-based retribution and rehabilitation.

A final caveat is worth mentioning: this Article uses scientific insights into moral judgments and decision-making to provide a theoretical model of legal insanity. Its aim is to interpret knowledge emerging from said scientific data, and combine that knowledge with legal arguments that might lead to a theoretical reconsideration of traditional approaches to the insanity defense. Considering its doctrinal scope, this Article does not address several practical issues concerning the implementation of this alternative model of legal insanity in forensic settings. These issues range from how the newly introduced emotional prong of insanity standards

Adolphs, *How Should Neuroscience Study Emotions? By Distinguishing Emotion States, Concepts, and Experiences*, 12 SOC. COGN. & AFFECTIVE NEUROSCI. 24, 27 (2017) (suggesting that emotions are biological functional states that regulate behavior and allow us to cope with environmental challenges. However, emotion states are not the same as conscious experiences of emotion (i.e., feelings). Rather, emotion states cause the conscious experience of emotions (i.e., feelings). Put this way, feelings are *derivative* of emotion states).

should be assessed through reliable judgments (both empirical and normative) to how the criminal justice system should appropriately deal with contentious classes of offenders who could be eligible for the new insanity defense. These issues are critical, and they will be explored thoroughly in future works.

I. THE RATIONALISM BEHIND THE INSANITY DOCTRINE

The insanity defense is an affirmative defense whereby criminal defendants seek to be excused from criminal liability on the grounds that, at the time of the crime, a mental illness deprived them of their relevant capacities required for criminal responsibility.⁹ In theoretical terms, insanity is a legal concept, not a medical one.¹⁰ While a mental illness (disease, defect, or disorder, depending on which terminology is adopted) in the clinical sense generally constitutes the “but for” condition of legal insanity, it must ultimately satisfy predetermined legal criteria to rise to the status of insanity.¹¹ The law is therefore never really interested in mental illness as such. There must certainly be mental illness, but there is always a second requirement, namely that the illness be of such form or degree that it meets certain legal criteria.¹² In sum, it is not a mental illness per se that provides grounds for excuse.¹³ Rather, the determination rests on whether

⁹ See Stephen J. Morse & Richard Bonnie, *Abolition of the Insanity Defense Violates Due Process*, 41 J. AM. ACAD. PSYCHIATRY & L. 488, 489 (“Blame and punishment by the state are fundamentally unfair and thus a violation of the Due Process Clause if an offender was not responsible for his crime. The affirmative defense of legal insanity applies this fundamental principle by excusing those mentally disordered offenders whose disorder deprived them of rational understanding of their conduct at the time of the crime.”).

¹⁰ See, e.g., Walter Sinnott-Armstrong & Ken Levy, *Insanity Defenses*, in THE OXFORD HANDBOOK OF PHILOSOPHY OF CRIMINAL LAW 299, 300 (John Deigh & David Dolinko eds., 2011) (observing that “[i]t is common to think that insanity is a medical condition. Psychiatrists, however, almost never describe their patients as ‘insane’ or ‘sane.’ . . . It is the judges and lawyers who have to decide who is insane and which mental conditions make someone insane. The law classifies some people as sane and others as insane in order to determine who should be held criminally responsible In this way, insanity is a legal concept.”).

¹¹ See, e.g., Paul Robinson, *The Effect of Mental Illness Under U.S. Criminal Law*, 65 N. IRL. LEGAL Q. 229, 230 (2014) (observing that “[i]t is not enough for the defense that an actor suffers from a mental disease or defect, even one that causes some dysfunction. To be held blameless, the actor’s mental illness must cause effects so strong that it would not be reasonable to expect the actor to have avoided the criminal law violation.”).

¹² HERBERT FINGARETTE, *THE MEANING OF CRIMINAL INSANITY* 45 (1972) (clarifying that “[w]hen criminal law asks questions concerning mental disease and insanity, it is concerned with the defendant’s mental capacities *with respect to the law*”).

¹³ Belief to the contrary risks giving rise to what Stephen Morse has defined as the “fundamental psycho-legal error.” See Morse, *supra* note 3.

the mental disease compromises the person's *capacity to be and act as a rational moral agent*.¹⁴

The capacity for moral rationality¹⁵ constitutes the benchmark of culpability and criminal responsibility.¹⁶ For agents to be considered responsible and therefore deserving of punishment, they must have the capacity to adopt a decision and make a choice against a system of moral and legal values.¹⁷ In fact, the fundamental assumption underlying the ideal of punishing only blameworthy agents is that culpable agents are practical

¹⁴ Helen Howard, *Diminished Responsibility, Culpability, and Moral Agency*, in *MENTAL CONDITION DEFENCES AND THE CRIMINAL JUSTICE SYSTEM* 318, 321 (Ben Livings, Alan Reed & Nicola Wake eds., 2015) ("Criminal responsibility will generally require a link to moral blameworthiness/culpability . . . Moral blameworthiness, in addition, presupposes that [an individual] is a rational moral agent who has sufficient understanding of his acts and deserves moral blame. Therefore, without moral agency there can be no culpability; without culpability there should be no criminal responsibility."); see also Peggy Sasso, *Criminal Responsibility in the Age of "Mind-Reading,"* 46 AM. CRIM. L. REV. 1191, 1193–94 (2009) ("[A]n individual who possesses those minimal capacities to qualify as a moral agent . . . is capable of engaging in conduct that rejects the community's moral norms.").

¹⁵ There are many definitions of rationality. Also, the meaning of rationality depends on the specific field of study, or context, where this concept is used. To avoid conceptual confusion, this Article only considers a legal notion of rationality, one which fits the purposes of criminal law. See Stephen J. Morse, *Rationality and Responsibility*, 74 S. CAL. L. REV. 251, 252 (2002) (describing [the capacity for moral] rationality as "the ability to perceive accurately, to get the facts right, to form justifiable beliefs, and to reason instrumentally, including weighing the facts appropriately and according to a minimally coherent preference-ordering. Rationality includes the general ability to recognize good reasons that should guide action. Put yet another way, it is the ability to act for good reasons."); see also Anthony Duff, *Answering for Crime*, 106 PROC. ARISTOTELIAN SOC'Y 87, 90 (2006) (likewise asserting that we are responsible agents "insofar as we are capable of grasping and being guided by reasons and of answering for ourselves in terms of reasons. Responsibility is in play when reasons are in play: we exercise our capacities for responsible agency in responding to reasons; we are responsible for such exercises, and for our failures to exercise those capacities when we fail to respond to reasons that bear on our thoughts and actions.").

¹⁶ See, e.g., Stephen J. Morse & Morris Hoffman, *The Uneasy Entente Between Legal Insanity and Mens Rea: Beyond Clark v. Arizona*, 97 J. CRIM. L. & CRIMINOLOGY 1071, 1117 (2007) ("Rationality is the touchstone of criminal responsibility, as the structure of criminal law itself indicates. All laws, criminal and civil, make sense and are functional precisely because they provide action-guiding reasons addressed to potentially rational creatures It is simply unfair to hold responsible . . . wrongdoers who . . . were not capable of being rational at the time of the crime.").

¹⁷ Stephen J. Morse, *Determinism and the Death of Folk Psychology: Two Challenges to Responsibility from Neuroscience*, 9 MINN. J.L. SCI. & TECH. 1, 5 (2008) ("The law properly treats persons generally as intentional creatures and not as mechanical forces of nature. Law and morality are action-guiding and could not guide people ex ante and ex post unless people could use rules as premises in their practical reasoning.").

moral reasoners;¹⁸ they are capable of reasoning instrumentally about facts as they relate to social, moral, or legal norms and of determining their conduct by virtue of normative considerations about what they ought or ought not to do.¹⁹

In the eyes of the law, the capacity for moral rationality—or capacity for practical moral reasoning—essentially includes two mental prongs: cognition and volition.²⁰ The cognitive prong of moral rationality generally regards the *capacity for understanding, or knowing, the factual and moral significance of a given action performed.*²¹ Thus, culpable agents are those

¹⁸ See, e.g., Douglas Husak, “Broad” Culpability and the Retributivist Dream, 9 OHIO ST. J. CRIM. L. 449, 465 (2012) (assuming without argument that reasons to which persons must respond in order to become eligible for blame and punishment are indeed *moral reasons*. According to Husak, responding to reasons does not solely mean “ability to conform to moral reasons,” but also to “understand the special motivating force of moral reasons.”).

¹⁹ See, e.g., Stephen J. Morse, *Reason, Result, and Criminal Responsibility*, U. ILL. L. REV. 363, 368 (2004) (arguing that “[l]egal and moral rules are not simply mechanistic causes that produce “reflex” compliance. They operate within the domain of practical reason. Agents are meant to and can only use these rules as potential reasons for action as they deliberate about what they should do. Moral and legal rules thus guide actions primarily because they provide an agent with good moral or prudential reasons for forbearance or action. Unless people were capable of understanding and then using legal rules as premises in deliberation, law would be powerless to affect human behavior.”).

²⁰ The inclusion of cognition and volition as the essential prongs of the capacity for moral rationality has received wide acceptance in various theories of criminal responsibility. For instance, Herbert Hart’s capacity-responsibility theory describes the relevant capacities necessary for criminal responsibility as “understanding, reasoning and controlling conduct: the ability to understand what conduct legal and moral rules require, to deliberate and reach decisions concerning these requirements; and to conform to decisions when made.” See H.L.A. HART, PUNISHMENT AND RESPONSIBILITY: ESSAYS IN THE PHILOSOPHY OF LAW 222 (1967). Endorsing a hybrid version of the character-based and choice-based theories of criminal responsibility, Antony Duff similarly holds that “someone who is to be held responsible for his choices must at least be capable of recognizing the relevant empirical aspects of his actions and its circumstances, and of foreseeing its consequences; he must also have the kind of ‘instrumental rationality’ which enables him to *determine* which actions will serve whatever ends he has.” See Robin Antony Duff, *Choice, Character, and Criminal Liability*, 12 L. & PHIL. 345, 356 (1993). Elsewhere, Duff argues that control too is an essential prerequisite of criminal responsibility. Control, according to Duff, “is a matter of rational capacities: thus I have control over my actions insofar as I have the capacities necessary to recognize reasons and guide my actions by them, insofar as I am capable of engaging in practical reasoning and of actualizing its results.” See Robin Anthony Duff, *Who is Responsible, for What, to Whom?*, 2 OHIO ST. J. CRIM. L. 441, 452 (2005).

²¹ HERBERT FINGARETTE & ANN FINGARETTE HASSE, MENTAL DISABILITIES AND CRIMINAL RESPONSIBILITY, 224–26 (1979) (claiming that “among members of a community there is a certain valuational ‘background nexus’ of basic perception and basic values. [Therefore] the basic factual presumption that underlies our expectation that people should deliberately [act lawfully] is that each individual in the community shares in a practical

who know or can understand what they are doing and the practical consequences of their actions, as well as the fact that their actions contradict society's morals.²² On the other hand, volition consists of the *free exercise of choice of a given course of action among conflicting reasons*.²³ Essentially, volition encompasses the capacity to exert self-control and resist impulses to engage in certain conduct.²⁴ It follows that culpable agents are those who, having conflicting reasons, *can* do otherwise (i.e., they are equipped with the capacity to resist their impulses), but simply *will not* do otherwise and thus choose to act upon their antisocial impulses.²⁵

Critically, the legal understanding of capacity for moral rationality is profoundly cognitivist. Embracing the rationalist view of rational thought and behavior—law-abiding and antisocial alike—endorsed by classical and neo-classical thinkers,²⁶ contemporary criminal law presupposes that the

awareness of this background-nexus of basic perceptions and basic values. [For instance] there is the person who has a *practical* grasp of the general moral significance of killing, and who has the sense that it is . . . an issue of deep concern to the law, and generally forbidden—but who may make a personal judgment that a particular unlawful killing is acceptable or even desirable. Such a person is rational in regard to the law, even though . . . in disagreement with it”).

²² Michael Corrado, *Responsibility and Control*, 34 HOFSTRA L. REV. 59, 65 (“Only those who have a reasonable awareness of the consequences of their actions . . . may be held responsible for what they do.”).

²³ MARK D. WHITE, THE MISTAKEN QUEST FOR A CONTROL TEST: FOR A RATIONALITY STANDARD OF SANITY 196 (2017) (“Control (or ‘self-control’) involves the ability to resist an immediate desire for the sake of one’s values or long-term goals. For a defendant to claim an incapacity for self-control is essentially to assert that, according to his better judgment, he did not truly want to commit the crime on the one hand, and his better judgment on the other. Without such a tension . . . then no reason exists to assert he experienced a defect of will.”).

²⁴ See Stephen J. Morse, *Culpability and Control*, 142 U. PA. L. REV. 1587, 1587 (1994) (“‘I couldn’t help myself’; ‘I had no choice’; ‘I couldn’t control myself’; ‘I was forced to do it.’ All are common explanations used to support the claim that an agent is not morally or criminally responsible for otherwise culpable conduct. The most common criminal law ‘control’ excuses that instantiate these claims are duress and the so-called ‘volitional’ tests for legal insanity.”).

²⁵ Rebecca Hollander-Blumoff, *Crime, Punishment, and the Psychology of Self-Control*, 61 EMORY L. J. 501, 515 (2012) (“[A] criminal is a person who deliberately chooses to engage in behavior that he knows is wrong; we call this crime because the person *could have* chosen not to engage in the behavior. That is, that person could have exercised self-control over his actions and opted for a different choice.”).

²⁶ See J. M. Canals, *Classicism, Positivism, and Social Defense*, 50 J. CRIM. L. & CRIMINOLOGY 541, 543 (1960); Clarence Jay Jeffrey, *The Historical Development of Criminology*, 50 J. CRIM. L. & CRIMINOLOGY 3 (1959) (discussing generally the main differences between Classical and Neo-classical schools, and the Positive school of criminal law and criminology); Raed SA Faqir, *The Philosophy of Punishment: A Study to the History*

capacity for moral rationality is entirely governed by the powers of “reason,” namely the active area of the mind²⁷ that is usually identified with cognitive faculties like *reasoning, thinking, planning, learning, understanding*, and the like.²⁸ Thus, contemporary criminal law identifies “reason”—*qua* the only source of the capacity for moral rationality—with the cognitive dimension of the mind. As a synonym for reason and the highest function of the mind, cognition (and its related faculties) is what drives conscious decision-making, intention-forming, and planning. It is also what guides individuals’ practical moral reasoning about what they ought and ought not to do.²⁹ Cognition is thus the sole component of the legally relevant mind.

Put yet another way, contemporary criminal law grounds the notion of the capacity for moral rationality in rationalist principles to espouse a cognitive model of culpability. As long as individuals display intact cognitive faculties, they are presumed to have the capacity to reason instrumentally about the factual and moral consequences of their conduct,

of Classical and Positive Schools of Penology, 1 FORENSIC RES. CRIMINOL. INT. J., 1, 3-5 (2015); KATHERINE WILLIAMS, TEXTBOOK OF CRIMINOLOGY 7–18 (2012).

²⁷ Christine M. Korsgaard, *The Activity of Reason*, 83 PROC. & ADDRESSES AM. PHIL. ASS’N 23, 30 (2009) (“The faculty of reason is not identified merely as the ability to recognize and respond to reasons. The faculty of reason is identified rather as the active dimension of the mind, and rational principles are then identified as those that describe or constitute rational activity.”); *see also* CARLSON ANYANGWE, CRIMINAL LAW: THE GENERAL PART 249 (2015) (reporting that “[t]he cognitive functions are our intellectual functions, that is, our ability to assimilate information from the environment, remember it, organise and process it in a rational manner, drawing rational conclusions, and making appropriate decisions. Thus, they include our functions of perception, thinking, reasoning, understanding, judgement and recall.”).

²⁸ FINGARETTE, *supra* note 12 at 181–82 (describing reason as “that guiding or directing faculty of the mind . . . by virtue of which . . . man has traditionally been said to be a ‘rational being.’” Fingarette holds that “the concept of reason has . . . been taken as the key . . . to man’s thinking, to what are called his cognitive capacities”).

²⁹ *See, e.g.*, MICHAEL MOORE, LAW AND PSYCHIATRY: RETHINKING THE RELATIONSHIP 83 (1984) (discussing the relationship between wrongdoing and culpability in the context of negligence, he observes that the negligent actor “is . . . capable of calculating what actions are likely to lead to what results and even to assign relative probabilities to each. He is, in other words, a pre-eminent practical reasoner, finding the morally and legally correct major premises . . . and forming the accurate means/end beliefs . . . for his minor premises It is because people have the capacity to reason this way that they can be said to be culpable when they do not do so.” He further explains that “[the] failure to make the right cost/benefit calculation” makes people culpable only if they have the capacity to reason this way); *see also* ANYANGWE *supra* note 27 at 249 (“[a] person’s ability to distinguish between right and wrong; to know when certain conduct would be wrong according to the standards of law, society and/or morality; and to decide on an appropriate course of action, depends on that person’s cognitive functions.”).

as well as to choose to engage in a given conduct—either right or wrong—rather than in another.³⁰ As such, they are assumed to be fully accountable for their behavior.

The preeminent role of cognition (and of cognitive faculties) as the absolute component of the capacity for moral rationality is markedly reflected in insanity standards. Insanity tests, as is incontrovertibly evidenced by the history of conceptions of insanity, are mostly cognition based.³¹ The reason is easy to grasp: if culpability requires that people possess a sufficient degree of intellectual capacity to reason instrumentally about the factual and moral consequences of engaging in criminal conduct and determine their actions accordingly, a lack of culpability can be found only when these conditions are not met. The sub-sections that follow explore these claims in more detail. They do so through an analysis of each prong—the cognitive and the volitional—in the most popular insanity tests that have been adopted in the United States, namely the *M’Naghten* rule and the American Law Institute (ALI) test. The purpose is not to conduct an in-depth analysis of these insanity standards (the literature is already overabundant),³² but simply to highlight the main normative aspects that will allow one to deduce the law’s acceptance of an intellectualistic understanding of the capacity for moral rationality underlying the insanity doctrine.

A. COGNITION AND THE COGNITIVE PRONG OF INSANITY TESTS

The strict relationship between capacity for moral rationality and cognition can be easily grasped in the cognitive prong of insanity standards. To illustrate this relationship, this section analyzes first the formulation of

³⁰ Glanville Williams, *The Criminal Responsibility of Children*, CRIM. L. REV. 493, 494 (1954) (asserting that “the only persons capable of acting wrongly are those of a certain intelligence or intellectual accomplishment”).

³¹ See *Clark v. Arizona*, 548 U.S. 735, 750 n.12 (2006) (listing state statutes); *infra* Section I.A and Section I.B.

³² See generally Michelle Holtzman, *Criminal Insanity – Another M’Naghten?*, 23 U. MIAMI L. REV. 644 (1969); American Bar Association, *Insanity Defense*, 16 MENTAL & PHYSICAL DISABILITY L. REP. 27 (1992); JANE CAMPBELL MORIARTY, *THE ROLE OF MENTAL ILLNESS IN CRIMINAL TRIALS: THE INSANITY DEFENSE 2* (2001); Lisa A. Callahan et al., *The Volume and Characteristics of Insanity Defense Pleas: An Eight-State Study*, 19 BULL. AM. ACAD. PSYCHIATRY & L. 331 (1991); GERBEN MEYNEN, *LEGAL INSANITY EXPLORATIONS IN PSYCHIATRY, LAW AND ETHICS* (2016); Susan Rozelle, *Pure Insanity*, 42 TEX. TECH. L. REV. 543 (2009); Paul Robinson & Markus Dubber, *The American Model Penal Code: A Brief Overview*, 10 NEW CRIM. L. REV. 319 (2007); Norval Morris, Richard Bonnie & Joel Finer, *Should the Insanity Defense Be Abolished? An Introduction to the Debate*, 1 J.L. & HEALTH 113 (1986–1987).

the *M'Naghten* rule.³³ The test contained in it states that insanity exists when the following conditions are met:

[A]t the time of the committing of the act, the party accused was laboring under such a defect of reason, from disease of [the] mind, as not to know the nature and quality of the act he was doing, or, if he did know it, that he did not know he was doing what was wrong.³⁴

The cognitivism upon which the rule relies emerges from the relationship among “disease of [the] mind,” “defect of reason,” and “knowledge.” The concept of “reason” in the formulation of the *M'Naghten* rule has not been fully explored by criminal law theorists, and its meaning remains unclear.³⁵ According to Herbert Fingarette, however, the expression “defect of reason” as it is used in *M'Naghten* clearly refers to a defect in the capacity for (moral) rationality.³⁶ He argues that the concept of rationality has, for the most part, fallen by the wayside during debates on the insanity defense, since the *M'Naghten* test’s “defect of reason” phrase “has not been understood.”³⁷ Fingarette holds that this is a “profound mistake [because] the defect-of-reason clause tells us that ‘know the nature and quality of the act’ and ‘know that is wrong’ must be taken to apply with reference to the person’s . . . capacity for rational conduct.”³⁸

Accepting this interpretation of the defect of reason as defect of the capacity for moral rationality, the intellectualistic substance of this clause emerges primarily from the knowledge requirement. In this respect, Arval Morris emphasizes the pivotal function of the word “know,” for it “circumscribes the entire test by singling out one aspect of a human being’s total personality, the cognitive one.”³⁹ Therefore, Morris continues, “the

³³ The *M'Naghten* rule has been largely qualified as the product of rationalist psychology in vogue at the time it was enacted. See, e.g., Note, *Criminal Responsibility and Proposed Revisions of the M'Naghten Rule*, 32 ST. JOHN'S L. REV. 247, 250 (1958) (“*M'Naghten*, it is claimed, is a product of a rationalist era, acknowledging only the cognitive or intellectual faculty and does not allow for the incapacity of the will or the influence of the emotions.”); see also RUDOLPH JOSEPH GERBER, *THE INSANITY DEFENSE* 30 (1984) (asserting that at the time of *M'Naghten*, “cognition was seen as the highest function of the personality. Philosophers searching for the Cartesian dregs of the period expressed the notion that the mind controlled bodily behavior like an angel driving a machine.”).

³⁴ R. v. *M'Naghten*, 8 Eng. Rep. 718 (1843).

³⁵ See Walter Sinnott-Armstrong & Ken Levy, *Insanity Defenses*, in *THE HANDBOOK OF PHILOSOPHY OF CRIMINAL LAW* 299, 306 (John Deigh & David Dolinko eds., 2011).

³⁶ FINGARETTE, *supra* note 12, at 198 (“I believe that ‘a defect of reason’ from ‘disease of the mind’ is to be read in paraphrase: ‘substantial defect in capacity for rational conduct’ as ‘an endogenous (pathological) condition of mind.’”).

³⁷ *Id.*

³⁸ *Id.*

³⁹ Arval A. Morris, *Criminal Insanity*, 43 WASH. L. REV. 583, 605 (1968).

test is heavily intellectualistic, and from a psychological point of view, narrow because the cognitive becomes the single, important criterion of criminal responsibility.”⁴⁰ Likewise, Rudolph Gerber critically observes that the term “knowledge” traditionally refers to a verbal or purely intellectual assent to a moral proposition.⁴¹ The word “know” can therefore be interpreted as an appreciation of “the significance of cognitive observation, that is, whether the defendant is able to relate what is known to the situation at hand and to govern conduct accordingly.”⁴² In this sense, the verb “to know” indicates one’s capacity to be aware of and correctly understand certain objective features of behavior. What is more, as Raider notes, courts tend not to define the verb “to know,” and therefore its interpretation is left largely to juries’ discretionary common sense.⁴³

The knowledge requirement circumscribes the meaning of the “disease of the mind” clause to encompass only cognitive diseases. As has been observed, courts vary on how they define “disease of [the] mind.”⁴⁴ Hence, a finding of such disease “follows almost automatically”⁴⁵ when it is found that a defendant was in such a state that he or she did not know the factual and moral implications of his or her conduct.⁴⁶ However, an explicit interpretation of the substance of the “disease of the mind” requirement can be found in *R v. Kemp*,⁴⁷ a case that was heard at the Bristol Assizes in 1957. Commenting on the relationship between the “defect of reason” and “disease of the mind” requirements, Justice Devlin wrote that “[t]he law is not concerned with the brain but with the mind, *in the sense that ‘mind’ is*

⁴⁰ *Id.*

⁴¹ Rudolph Joseph Gerber, *Is the Insanity Test Insane?*, 20 AM. J. JUR. 111, 120 (1975).

⁴² RICHARD BONNIE ET AL., A CASE STUDY IN THE INSANITY DEFENSE: THE TRIAL OF JOHN HINCKLEY, JR. 12 (3d ed. 2008); *see also* ABRAHAM GOLDSTEIN, THE INSANITY DEFENSE 49–50 (1967).

⁴³ Laura Raider, *Toward a New Test for Insanity Defense: Incorporating the Discoveries of Neuroscience into Moral and Legal Theories*, 46 UCLA L. REV. 289, 306 (1998); *see also* James R.P. Ogloff, *A Comparison of Insanity Defense Standards on Juror Decision Making*, 15 L. & HUM. BEHAV. 509, 526 (1991); NORMAN FINKEL, COMMONSENSE JUSTICE: JURORS’ NOTIONS OF THE LAW (1995).

⁴⁴ JOHN PARRY, CRIMINAL MENTAL HEALTH AND DISABILITY LAW, EVIDENCE, AND TESTIMONY 336 (2009) (“Under *M’Naghten*, there must be a ‘disease of the mind.’ This term has been interpreted very differently over the years. In its broadest meaning, ‘any diagnosable mental disorder’ is sufficient, while under the more narrow interpretation a ‘severe impairment, usually in the form of psychosis,’ is required.”).

⁴⁵ Richard H. Kuh, *The Insanity Defense—An Effort to Combine Law and Reason*, 110 U. PA. L. REV. 771, 785 (1962).

⁴⁶ *Id.*

⁴⁷ 1 QB 399 (1957). For a comment, *see* J.E. Hall Williams, *Defect of Reason from Disease of the Mind*, 20 MOD. L. REV. 55, 56 (1957).

ordinarily used, the mental faculties of reason, memory and understanding."⁴⁸

This interpretation of the "disease of mind" requirement is pivotal because it underlines the equation of the (legally relevant) mind with cognitive faculties.⁴⁹ Hence, because the rule states that the defect of reason (i.e., moral rationality) must derive from a disease of the mind, it implicitly assumes that the only mental source of the capacity for moral rationality is cognition. Consequently, only purely cognitive defects are considered capable of giving rise to irrational behavior. As Gerber more precisely describes it, it is only "when cognition is defective the personality as a whole is so impaired that the accused cannot 'know' the wrongfulness of his actions."⁵⁰ Thus, a legally relevant disease of the mind cannot but be a cognitive disease, such as pure psychosis.

The test contained in the *M'Naghten* rule has given rise to several controversies over the years. On the one hand, some specialists in criminal law have been staunch defenders of the test and maintained that it is fundamental to the notions of moral blame and retribution.⁵¹ On the other hand, mental health specialists, supported by some legal scholars and practitioners, have held that the *M'Naghten* test is obsolete and unscientific⁵² and "not only bad science but unsound law."⁵³ Critics holding this view have relied upon the fact that cognitive or intellectual integrity is not sufficient to account for a person's knowledge of the wrongfulness of the act he or she is about to perform and the ability to control himself or herself.⁵⁴ In other words, the intellectual test contained in the rule is too narrow to encompass the entire scope of legal insanity.

⁴⁸ *R v. Kemp*, 1 QB at 399 (emphasis added).

⁴⁹ See FINGARETTE, *supra* note 12, at 144 ("[T]he *M'Naghten* test deals with only one of the mind's three 'functions,' the cognitive function.").

⁵⁰ Gerber, *supra* note 41, at 119.

⁵¹ See, e.g., Jerome Hall, *Responsibility and Law: In Defense of the McNaghten Rules*, 42 A.B.A. J. 917 (1956).

⁵² See Simon E. Sobeloff, *Insanity and the Criminal Law: From M'Naghten to Durham, and Beyond*, 41 A.B.A. J. 793, 877 (1955) ("The *M'Naghten* rule requires medical witnesses to testify in terms that to them are artificial and confining . . . When [a doctor] is forced to adopt the vocabulary of morality and ethics, he is speaking in what to him is a foreign language . . .").

⁵³ Kuh, *supra* note 45, at 782.

⁵⁴ See, e.g., Carl Cohen, *Criminal Responsibility and the Knowledge of Right and Wrong*, 14 U. MIAMI L. REV. 30, 43 (1959) ("The emphasis upon knowledge as the test of responsibility is indicative of what is perhaps the underlying misconception of the *McNaghten* rule—that the cognitive capacities can be singled out, among mental phenomena, as the *proper* determinants of sanity and responsibility. Although we do, for purposes of analysis, distinguish the cognitive, conative, and affective aspects of mental life,

To fill the gaps and shortcomings of the *M'Naghten* rule, in 1962 the ALI introduced a new insanity standard in the MPC.⁵⁵ The MPC's formulation is based on the assumption that insanity should be grounded in a broader understanding of cognition, and a reference to volitional incapacity should be included explicitly in the formulation of the defense.⁵⁶ The broadening of the cognitive prong of legal insanity has resulted in the dismissal of the verb "to know" and the simultaneous adoption of the broader verb "to appreciate."⁵⁷ This term has, however, caused interpretive confusion because it is unclear how broadly it should be understood.⁵⁸ According to some authors, the term "appreciate" has been introduced to allow the inclusion of the emotional capacity to perceive the wrongfulness of one's conduct.⁵⁹ Others have held that this broader understanding of the verb "to appreciate" is merely ostensible, and that it still refers to a purely cognitive capacity to propositionally understand the meaning of one's conduct.⁶⁰

This stricter cognition-based interpretation is confirmed by the so-called "caveat paragraph" of the test, which explicitly rules out "any abnormality manifested only by repeated criminal or otherwise antisocial

it is an error to assume, as the *McNaghten* rule does, that these aspects of experience can be separated to the extent that one alone—the cognitive—is the index of mental health. The actual mental experience of persons, well or sick, cannot be broken up so neatly into its constituents. We can no longer assume that reason, or cognition, is the only—or even the prime—regulator of conduct. That the human personality is an integrated unity, in the direction of which all of the modes of experience play some part is now a commonplace. The point is that people may be brought, by insane emotion or compulsion, to do what they themselves know to be wrong. Yet that faculty psychology already discussed, so prevalent in the nineteenth century, has colored, through the *McNaghten* opinions, the legal standards of the present day.”).

⁵⁵ MODEL PENAL CODE § 4.01 (AM. LAW INST. 1962).

⁵⁶ See generally MODEL PENAL CODE (AM. LAW INST. 1985).

⁵⁷ MODEL PENAL CODE § 4.01(1) (AM. LAW INST. 1962) (“A person is not responsible for criminal conduct if at the time of such conduct as a result of mental disease or defect he lacks substantial capacity . . . to appreciate the criminality of his conduct.”).

⁵⁸ See Kuh, *supra* note 45, at 797–98 (asserting that the words ‘substantial’ and ‘appreciate’ “were intentionally chosen for their imprecision”).

⁵⁹ RITA J. SIMON & DAVID E. AARONSON, *THE INSANITY DEFENSE: A CRITICAL ASSESSMENT OF LAW AND POLICY IN THE POST-HINCKLEY ERA* 37–39 (1988).

⁶⁰ See GOLDSTEIN, *supra* note 42, at 88. Also, the cognitive nature of the verb “to appreciate” emerges from the explanatory notes of Section 4.01 of the Model Penal Code: “An individual’s failure to appreciate the criminality of his conduct may consist in a lack of awareness of what he is doing or a misapprehension of material circumstances, or a failure to apprehend the significance of his actions in some deeper sense.” THE AMERICAN LAW INSTITUTE, MODEL PENAL CODE OFFICIAL DRAFT AND EXPLANATORY NOTES, PART I. GENERAL PROVISIONS 62 (1985).

conduct from the notion of mental disease or defect.”⁶¹ The reference, as has been claimed, is to those categories of subjects with impaired emotional faculties and marked tendencies to engage in antisocial conduct, such as in the case of psychopaths.⁶²

B. COGNITION AND THE VOLITIONAL PRONG OF INSANITY TESTS

The volitional prong of insanity standards is surely more controversial than the cognitive one. As noted above, volition essentially means free exercise of choice among conflicting courses of action, with distinct sets of reasons.⁶³ In substantial capacity doctrines, such as insanity, volition fundamentally equates to the capacity for self-control.⁶⁴ Unlike the cognitive prong, however, the volitional prong of insanity standards is both conceptually and epistemologically troublesome.

Conceptually, it is not clear what the capacity or incapacity for self-control actually means, nor is there a unanimous consensus about it among philosophers, legal scholars, and legal practitioners. Unlike cognitive capacity or incapacity, volitional capacity or incapacity results in a doubling of a person’s personality, meaning that normality and abnormality coexist in the same person, between which it is incredibly difficult to conceptualize a dividing line.⁶⁵ In other words, it is not clear when a person *cannot* resist an impulse or simply *will not* resist an impulse. As a consequence, from an epistemological perspective, the major reasons for the skepticism surrounding the issue of whether lack of self-control should be taken into account as an autonomous prong of insanity range from difficulty in assessing and quantifying self-control, to practical difficulty in differentiating lack of self-control from poorly planned or impulsive acts.⁶⁶

Based on these limitations, many scholars, and the law itself, endorse the view that the capacity for self-control ultimately depends on the cognitive faculties of knowledge or understanding.⁶⁷ In other words, if the

⁶¹ MODEL PENAL CODE § 4.01(2) (AM. LAW INST. 1962).

⁶² See RALPH SLOVENKO, *PSYCHIATRY IN LAW/LAW IN PSYCHIATRY* 197 (2d ed. 2009).

⁶³ See *supra* notes 23–24.

⁶⁴ See *supra* note 24.

⁶⁵ A remarkable attempt has been made by Michael Moore, who has provided a folk-psychological account of volitional capacity and excuse. See Michael S. Moore, *The Neuroscience of Volitional Excuse*, in *PHILOSOPHICAL FOUNDATIONS OF LAW AND NEUROSCIENCE* 179 (Dennis Patterson & Michael S. Pardo eds., 2016).

⁶⁶ GARY B. MELTON ET AL., *PSYCHOLOGICAL EVALUATIONS FOR THE COURTS: A HANDBOOK FOR MENTAL HEALTH PROFESSIONALS AND LAWYERS* 216 (3d ed. 2007).

⁶⁷ See, e.g., Jerome Hall, *Mental Disease and Criminal Responsibility - M’Naghten Versus Durham and the American Law Institute’s Tentative Draft*, 33 *IND. L.J.* 212, 213–14 (1958) (“The second point that I think we must hold on to is the relationship between

capacity for moral rationality means the ability to determine (i.e., know, be aware of, understand, and so forth) the factual, social, and moral value of a given action, it simultaneously consists of the ability “not to act [and, thus, to exert self-control] if doing so is known to be wrong.”⁶⁸ Thus, by presuming that the capacity for self-control depends on people’s prior knowledge of the factual and moral significance of their actions, criminal law implicitly accepts that the source of self-control lies in cognitive faculties.⁶⁹

The view of self-control as a faculty depending on knowledge—which is expressed in the ancient legal maxim *nihil volitum nisi praecognitum* (nothing is desired unless it is first known)—is manifest in the history of insanity standards. To begin, the *M’Naghten* test includes only a cognitive prong in that it roots the grounds for insanity in the possession of the cognitive capacity to know the nature and quality of one’s act and that the act is wrong.⁷⁰ As Gerber observes, “The [*M’Naghten*] rule assumes that if an individual ‘knows’ right from wrong, his rational powers are intact and that he is, therefore, capable of governing his conduct.”⁷¹ Similarly, Snyder claims that “an offender who knows what he does, knows that it is wrong . . . and coolly and carefully prepares what he does, can and does control his action right up to the moment of commission.”⁷² Thus, although a volitional prong is not explicitly provided, it is implicitly derived from the cognitive one.

intelligence and the control of conduct. If we look about us and visualize the magnificent structures of science and legal systems and ethics, we attribute these great achievements to man’s capacity for thought, to human understanding. Can we then allow psychiatrists or any other specialists to persuade us that human understanding has no effective relationship to the commission of the serious harms that are the concern of criminal law? It seems to me that we should ask for evidence and a great deal of evidence before we accept the irrationalism that one’s reason may be unimpaired and that nonetheless it exercises no control over such conduct.”).

⁶⁸ JEROME HALL, *GENERAL PRINCIPLES OF CRIMINAL LAW* 167 (2d ed. 1960). To put it into an example: if A knows that shooting B will kill B, and A knows that the act of killing is wrong, then A is supposed to refrain from shooting B because of this knowledge/awareness, both factual and normative.

⁶⁹ To put it into an example: A is intelligent, A’s intelligence enables him to know what stabbing B means, and therefore A willfully chooses to stab, or prevents himself from stabbing B, based on this knowledge.

⁷⁰ *R. v. M’Naghten* 8 Eng. Rep. 718 (1843).

⁷¹ GERBER, *supra* note 33, at 30.

⁷² Orvill C. Snyder, *Criminal Responsibility*, 1962 DUKE L.J. 204, 209 (1962); *see also* Kuh, *supra* note 45, at 782 (“As man is an ‘integrated personality,’ his knowledge, his will, and his ability to act are all intertwined. The word ‘know,’ as used in *McNaughton*, can be taken to mean not only the ability to perceive by use of the senses and intellect, but the ability to guide or control one’s action in the light of this perception.”).

As has also been noted, the *M’Naghten* rule has been harshly criticized by scholars and practitioners of criminal law who have emphasized the extreme narrowness of the sole cognitive test.⁷³ These critiques have led some United States jurisdictions, in an effort to counter the excessive restrictiveness of the *M’Naghten* rule, to adopt the “irresistible impulse test”⁷⁴—today abandoned—and, many years later, the ALI test.⁷⁵ Focusing solely on the latter, the ALI standard defines the capacity for self-control as “the capacity to conform behavior to what the law requires.”⁷⁶ Under this new definition, lack of self-control occurs when individuals cannot (or find it impossibly difficult) act as the law requires, regardless of whether or not they know they should behave as the law prescribes.

Despite its initial popularity, today only a few states continue to apply the ALI test.⁷⁷ Public outcry⁷⁸ after the verdict in *United States v. Hinckley*⁷⁹ led to subsequent legislation that narrowed the insanity defense by removing the volitional defense theory.⁸⁰ In 1983, the American Psychiatric Association released a statement arguing that volitional tests may be unnecessary because defendants who meet the exculpatory criteria set forth in volitional tests will usually meet the exculpatory criteria for cognitive impairment tests as well.⁸¹ In the wake of that statement, as of

⁷³ See *supra* notes 52–54 and accompanying text.

⁷⁴ See *Davis v. United States*, 165 U.S. 373, 378 (1897); *Parsons v. State*, 2 So. 854, 866–67 (Ala. 1877); *Commonwealth v. Rogers*, 48 Mass. (7 Met.) 500, 502 (1844).

⁷⁵ MODEL PENAL CODE (AM. LAW INST. 1985).

⁷⁶ MODEL PENAL CODE § 4.01 (AM. LAW INST. 1962).

⁷⁷ See American Academy of Psychiatry and Law, *AAPL Practice Guideline for Forensic Psychiatric Evaluation of Defendants Raising the Insanity Defense*, 42 J. AM. ACAD. PSYCHIATRY L. S3, S66 (2014).

⁷⁸ The Associated Press, *Hinckley Acquittal Brings Moves to Change Insanity Defense*, N.Y. TIMES, June 24, 1982, at D21.

⁷⁹ 525 F. Supp. 1342 (D.C. Cir. 1981).

⁸⁰ CHARLES PATRICK EWING & JOSEPH T. MCCANN, MINDS ON TRIAL: GREAT CASES IN LAW AND PSYCHOLOGY 99 (2006) (“[F]ederal and state legislators were busy revising or abolishing insanity laws in an effort to make sure that another Hinckley verdict would never occur.”).

⁸¹ American Psychiatric Association, *Statement on Insanity Defense*, 140 AM. J. PSYCHIATRY 681, 685 (1983) (“The line between an irresistible impulse and an impulse not resisted is probably no sharper than that between twilight and dusk The concept of volition is the subject of some disagreement among psychiatrists. Many psychiatrists therefore believe that psychiatric testimony (particularly that of a conclusory nature) about volition is more likely to produce confusion for jurors than is psychiatric testimony relevant to a defendant’s appreciation or understanding.”). However, the APA retired this position in 2008. See Donna Norris, *Reports to Membership*, 165 AM. J. PSYCHIATRY 1214, 1214 (2008) (“The sixth position statement, Insanity Defense, replaced the statement approved in 1982 and updated it with a more concise and up-to-date formulation. This more concise position

1984 the ALI test was largely discarded in favor of the Insanity Defense Reform Act (IDRA):⁸² the first federal codification of the insanity defense that was introduced with the explicit purpose of removing the volitional component from the ALI standard.

Today, a minor number of US jurisdictions employ the “control” test.⁸³ The majority of them remain mired in a view that the only viable basis for a plea of insanity is a mental disease or disorder that has led to cognitive impairment resulting in an incapacity to know, understand, or appreciate the factual, moral, social, or even legal significance of one’s conduct.⁸⁴ Lack of self-control is assessed indirectly as a possible further consequence of cognitive defects.⁸⁵ Therefore, if individuals are cognitively able to know right from wrong and their rational powers are

statement is aimed at underscoring APA support for a meaningful insanity defense without endorsing any particular standard.”).

⁸² 18 U.S.C. § 17 (1984) (“Affirmative defense: It is an affirmative defense to a prosecution under any Federal Statute that, at the time of the commission of the acts constituting the offense, the defendant, as a result of a severe mental disease or defect, was unable to appreciate the nature and quality or the wrongfulness of his acts. Mental disease or defect does not otherwise constitute a defense.”); see Stephen J. Morse, *Insanity Defense Reform Act (IDRA)*, in *ENCYCLOPEDIA OF PSYCHOLOGY AND LAW* 374 (Brian L. Cutler ed., 2008).

⁸³ See Paul H. Robinson & Tyler Scott Williams, *Mapping American Criminal Law: Variations Across the 50 States - Ch. 14 Insanity Defense*, U. PA. FACULTY SCHOLARSHIP 1718 (2017) [http://scholarship.law.upenn.edu/faculty_scholarship/1718].

⁸⁴ *Id.*

⁸⁵ See Paul Litton, *Is Psychological Research on Self-Control Relevant to Criminal Law?*, 11 OHIO ST. J. CRIM. L. 725, 730–31 (2014) (commenting on the relationship between the psychological conception of self-control and the understanding of self-control that is implicit in cognitive tests, he observes that “[a]n agent has the requisite control over her conduct, and is thus sane, if she has the capacity to understand the nature and moral quality of her conduct.” Using the famous example of Andrea Yates—who, while suffering from psychotic delusions, killed her five children—he observes that “[if we] stipulate that Yates had a very strong desire not to kill her children and that she fought against this desire because she believed that the morally best action was to kill her children . . . then Yates would have exercised self-control; however, she could still be judged insane under *M’Naghten* if she did not know her acts were wrong. The conception of control implicit in cognitive insanity standards is distinct from the kind of self-control under [psychological] study.”). See also *Yates v. State*, 171 S.W.3d 215 (Tex. App. 2005); for comments on Yates case and the relationship between cognitive and volitional prongs, see generally Brian D. Shannon, *The Time is Right to Revise the Texas Insanity Defense: An Essay*, 39 TEXAS TECH. L. REV. 67 (2006); Christine Michalopoulos, Note, *Filling in the Holes of the Insanity Defense: The Andrea Yates Case and the Need for a Volitional Prong*, 10 VA. J. SOC. POL’Y & L. 383 (2003); Melinda Carrido, *Revisiting the Insanity Defense: A Case for Resurrecting the Volitional Prong of the Insanity Defense in Light of Neuroscientific Advances*, 41 SW. L. REV. 309, 322–23 (2012).

intact, then it follows that they are capable of controlling their own conduct. Altogether, they do not meet the criteria to be eligible for legal insanity.

II. THE NEGLIGIBLE ROLE OF EMOTIONS

Criminal law's adherence to a cognitivist model of legal insanity finds further support in the lack of consideration of emotions within the essential components of the capacity for moral rationality—and hence of culpability.⁸⁶ While there is a long-standing dispute among scholars from different disciplines about the exact meaning of emotion,⁸⁷ this Article addresses an understanding of emotion adopted by criminal law. The answer is intuitive. Criminal law's view of emotion is rationalist and folk psychological.⁸⁸ Under the folk—and rationalist—conception, emotions mostly take on negative connotations, in that they are viewed as contrasting with intellect, and therefore may only undermine rational moral behavior.⁸⁹ Simply put, in criminal law “emotions are thought to be irrational, involuntary, and animal-like, whereas . . . intellect is rational, voluntary, and distinctly human.”⁹⁰

⁸⁶ As has been discussed, the legal understanding of the capacity for moral rationality is essentially based on the fusion of cognition and volition. *See supra* Part I.

⁸⁷ *See, e.g.*, Klaus R. Schreier, *What Are Emotions? How Can They Be Measured?*, 44 *Soc. Sci.* 695, 696 (2005) (“The concept of ‘emotion’ presents a particularly thorny problem. Even though the term is used very frequently, to the point of being extremely fashionable these days, the question ‘What is an emotion?’ rarely generates the same answer from different individuals, scientists or laymen alike.”); *see also* Ronald deSousa, *Emotion*, *Stanford Encyclopedia of Philosophy* (2017), <https://plato.stanford.edu/entries/emotion/> [<https://perma.cc/RVX9-RJXL>] (illustrating the different views about the functions and the ontology of emotions); John Deigh, *Concepts of Emotions in Modern Philosophy and Psychology*, in *THE OXFORD HANDBOOK OF PHILOSOPHY OF EMOTION* 17 (2009) (illustrating the different conceptions of emotions either as affective states or as cognitive states); Paul Kleinginna & Anne M. Kleinginna, *A Categorized List Of Emotion Definitions, With Suggestions For a Consensual Definition*, 5 *MOTIV. EMOT.* 345 (1981).

⁸⁸ Folk psychology can be defined as the natural tendency human beings have to express and describe the behavior of others on the basis of the possessing mental states (intentions, desires, beliefs, etc), by using common linguistic terms. Criminal law heavily embodies folk psychological accounts of human behavior. *See, e.g.*, Katrina Sifferd, *Translating Neuroscientific Evidence into the Language of the “Folk,”* in *NEUROSCIENCE AND LEGAL RESPONSIBILITY* 183, 191 (Nicole Vincent ed., 2010) (explaining that “criminal law grounds the assessment of responsibility on behavioral evidence, and behavioral evidence is likely to directly trigger attribution of mental states required by the legal criteria for guilt or not guilt.”).

⁸⁹ Commonplace idioms reveal this bias: “Keep a cool head;” “Keep your emotions at bay;” or “Do not let your passions interfere with your reason.”

⁹⁰ NORMAN FINKEL & GERROD PARROT, *EMOTIONS AND CULPABILITY: HOW THE LAW IS AT ODDS WITH PSYCHOLOGY, JURORS, AND ITSELF* 53 (2006).

This rationalist and commonsense-based understanding of emotions adopted by criminal law is what Dan Kahan and Martha Nussbaum call the “mechanistic conception”⁹¹ of emotions. Under the mechanistic conception, emotions have a negative effect on reasoning and self-control,⁹² as they are forces that do not contain or respond to thought.⁹³ As such, emotions are not part of the mental states that comprise moral rationality and do not play any positive role within morally rational decision-making. Rather, emotions are treated as irrational occurrences that may distort moral reasoning and potentially destabilize moral decision-making by preventing people from selecting the adequate means to achieve their goals, therefore negatively affecting their self-control.⁹⁴

While embracing this perspective, criminal law excludes the emotional dimension of the mind from having any significant relevance for the mental preconditions of culpability. As Norman Finkel and Gerrod Parrot note, “[t]he folk category of emotion can appear to threaten the orderly rule of law, for it carries with it the irrationality of primate impulses and the indeterminacy of subjective states. These perceived threats account for why the Law omits emotion in favor of more cognitive criteria. . . .”⁹⁵ Also, as the two authors suggest, the general reluctance to properly address emotion is linked to the law’s effort to avoid subjective rules, standards,

⁹¹ See Dan M. Kahan & Martha C. Nussbaum, *Two Conceptions of Emotion in Criminal Law*, 96 COLUM. L. REV. 269, 275–95 (1996).

⁹² See *id.* at 273.

⁹³ *Id.* at 278–79 (“[E]motions . . . are energies that impel the person to action, without embodying ways of thinking about or perceiving objects or situations in the world Emotions feel like things that sweep over us, or sweep us away, or invade us often without our consent or control . . .”).

⁹⁴ The other conception of emotion is the evaluative conception. See Kahan & Nussbaum, *supra* note 91, at 273 (explaining that under the evaluative conception “emotions express cognitive appraisals, that these appraisals can be morally evaluated, and that persons can and should shape their emotions through moral education”). Kahan and Nussbaum actually defend an idea that even though the law’s language may suggest a view of emotions that is distinctively primitive, bodily, and mechanical, its logic actually suggests an evaluative understanding of emotions. Speaking about rage, the two authors maintain that rage is mitigating not because of a mechanical loss of control but because rage expresses values that are proper in the situation. In other words, it is conceivable that the law expresses understanding of emotion’s cognitive aspects, even if it draws on non-cognitive metaphors to describe it. I disagree with this view. While the premises of Kahan’s and Nussbaum’s reasoning are logical and certainly correct (emotions *are* evaluative and cognitive, and *should be* considered as such by the law), I am not inclined to think that criminal law actually takes the evaluative perspective into account. As it emerges from normative texts and their dominant interpretations, the view that emotions are in opposition to cognition and self-control is manifest.

⁹⁵ See FINKEL & PARROT, *supra* note 90, at 48.

and judgments as much as possible.⁹⁶ Therefore, the law opts for “more mechanical, physical, bodily aspects of everyday emotion language.”⁹⁷

Despite criminal law’s firmness in denying emotion a positive role in a person’s capacity for moral rationality, some commentators do recognize that the presence or absence of emotional capacities or states should be weighed in the evaluation of the substantial capacities necessary for responsibility.⁹⁸ However, as Raider notes, “although some of these theorists acknowledge a richer conception of rationality than pure instrumental reasoning, including a limited role of emotions, they stop short of including all of the relevant capacities.”⁹⁹ Indeed, even when emotions are thought of as relevant components, they are still treated as a *quid pluris* and still on a distinct and minor level to that of cognition and rationality.¹⁰⁰

The neglected positive role of emotion in moral rationality is manifest in culpability doctrines and particularly in the insanity doctrine. The absence of emotions from the substance of legal insanity emerges from a twofold position: 1) no insanity standard provides for an emotional capacity test, that is, a test measuring a defendant’s capacity to emotionally appreciate the moral significance of his or her action; and 2) insanity standards infer volitional incapacity only from cognitive impairments, not emotional ones: emotional disturbance affecting self-control is considered only in diminished-capacity doctrines.

A. “MORAL INSANITY” AND THE LACK OF AN EMOTIONAL CAPACITY TEST

In 1835, the Bristolian physician Dr. James Cowles Prichard described moral insanity as “a form of mental derangement, a morbid perversion of the feelings, affections, and active powers without any illusion or erroneous conviction impressed upon the understanding: it sometimes co-exists with an apparently unimpaired state of intellectual faculties.”¹⁰¹ Prichard’s

⁹⁶ *Id.* at 83.

⁹⁷ *Id.*

⁹⁸ See MICHAEL MOORE, *PLACING BLAME: A GENERAL THEORY OF CRIMINAL LAW* 614–15 (1997); Stephen J. Morse, *Brain and Blame*, 84 *GEO. L.J.* 527, 543 (1996); Peter Arenella, *Convicting the Morally Blameless: Reassessing the Relationship between Legal and Moral Accountability*, 39 *UCLA L. REV.* 1511 (1992).

⁹⁹ Raider, *supra* note 43, at 293.

¹⁰⁰ For example, Stephen Morse qualifies emotions as one of the protective variables that also might help the agent to be in control. Yet, Morse places them on a distinct and independent level compared to rationality. See Morse, *supra* note 24, at 1607–08.

¹⁰¹ JAMES COWLES PRICHARD, *A TREATISE ON INSANITY AND OTHER DISORDERS AFFECTING THE MIND* 20 (1835). Also, Cesare Lombroso adhered to Prichard’s theory of moral insanity to complete his theory of the criminal man. Moral insanity, along with

description presents a view of people suffering from moral insanity as being impaired in their inclinations, tempers, habits, moral dispositions and natural impulses, but without any remarkable disorder or defect of the intellect or reasoning faculties, and particularly without any illusion or hallucination.¹⁰²

Presently, the concept of moral insanity has become synonymous with socio-affective disorders, such as antisocial personality disorder (ASPD), narcissistic personality disorder (NPD) and, above all, psychopathy.¹⁰³ Although appearing in different forms, these disorders share emotional deficits and an increased disregard for other individuals which can be expressed through a tendency to engage in violent and aggressive conduct.¹⁰⁴ People suffering from these disorders do not exhibit any significant deficit in the cognitive functions of knowledge and understanding.¹⁰⁵ Psychopathy represents the most emblematic example. Although there is a growing consensus that psychopathy should be classed as a mental disorder, criminal law treats psychopaths as *paradigmatic* culpable agents.¹⁰⁶ The reason is simple: psychopaths' intellectual faculties

atism and epilepsy, forms a universal trait of criminal subjects. See CESARE LOMBROSO, CRIMINAL MAN 188 (3d ed. 1884) transl. by Mary Gibson & Nicole Hahn Rafter, 2006).

¹⁰² PRICHARD *id.* at 16.

¹⁰³ According to the nosographic description contained in the DSM-5, Antisocial Personality Disorder is characterized by a pervasive pattern of disregard for the rights of other people, which often manifests as hostility, aggression, or both. Deceit and manipulation are also central features. People with Narcissistic Personality Disorder have significant problems with their sense of self-worth stemming from a powerful sense of entitlement. This leads them to believe that they deserve special treatment and to assume that they have special powers, are uniquely talented, or are especially brilliant or attractive. Their sense of entitlement can lead them to act in ways that fundamentally disregard and disrespect the worth of those around them. Psychopathy, on the other hand, is a mental disorder featured by marked emotional dysfunctions, limited capacity for moral judgments, and recidivistic offending. Psychopathy is not officially recognized as a personality disorder. It is not even included in the DSM-5, nor is there a unanimous opinion among psychiatrists and psychologists as to whether psychopathy should be qualified as a disorder at all. Attempts to define psychopathy as an autonomous kind of disorder and provide specific items to identify it have been made primarily by Dr. Robert Hare, who authored the Psychopathy Checklist Revised in 1990. As described in the checklist, psychopathy encompasses traits typical to both ASPD (e.g., lack of impulse control) and NPD (e.g., grandiose sense of self-worth). See Robert Hare et al., *The Revised Psychopathy Checklist: Reliability and Factor Structure*, 2 PSYCHOL. ASSESS.: J. CONSULT. CLIN. PSYCHOL. 338, 339 (Table 1) (1990).

¹⁰⁴ *Id.*

¹⁰⁵ *Id.*

¹⁰⁶ See Stephen J. Morse, *Psychopathy and Criminal Responsibility*, 8 NEUROETHICS 205, 207 (2008) ("The law does not excuse psychopaths, even those whose psychopathy is clear and severe. Psychopathy is not a legally sufficient basis to raise an insanity defense or any other excuse.").

are substantially intact; psychopaths are intelligent enough to know the facts, have no misperception of reality, and understand that there are rules and consequences for violating them.¹⁰⁷ What they lack is “simply” empathy, regret, guilt, or more generally, prosocial emotions and feelings.¹⁰⁸ Even if psychopaths display significant emotional abnormalities, they are never legally excused for their criminal actions.¹⁰⁹ Rather, their condition may even lead to harsher sentences.¹¹⁰

The first and fundamental reason for excluding subjects suffering from severe socio-affective deficits—such as psychopathy—from the range of eligible candidates for the insanity defense lies in the law’s limited regard of emotional faculties within the evaluation of the capacity for moral rationality.¹¹¹ As such, deficits in the emotional faculties—however pathological—are not considered to sufficiently affect an individual’s capacity for moral rationality.¹¹² As long as individuals possess intact intellectual faculties, and thus demonstrate substantive instrumental reasoning faculties and competent use of them, a possible deficiency in emotional faculties does not affect their culpability.¹¹³ Quite the contrary, these characteristics appear to symptomize an intense form of “evil” which deserves even harsher punishment.¹¹⁴

The second, and consequential, reason for excluding socio-affective deficits from the host of conditions eligible for the insanity defense is that they—however pathological—do not meet traditional insanity standards, which admit no relevance for emotional capacities. While standards such as the *M’Naghten* rule do not consider emotional capacity at all, the MPC seems to admit the lack of emotional capacity has some relevance.¹¹⁵ As noted above, however, this acceptance is ostensible.¹¹⁶

¹⁰⁷ *Id.* at 208.

¹⁰⁸ See Hare, *supra* note 103.

¹⁰⁹ See Morse, *supra* note 106, at 208.

¹¹⁰ *Id.* at 207–08.

¹¹¹ See *supra* Part II.

¹¹² See *supra* Part II.

¹¹³ See *supra* Part II.

¹¹⁴ See, e.g., John F. Edens et al., *Psychopathy and the Death Penalty: Can the Psychopathy Checklist-Revised Identify Offenders Who Represent ‘A Continuing Threat to Society’?*, 29 J. PSYCHIATRY & L. 433, 458–63 (2001); Stephen D. Hart, *Psychopathy, Culpability, and Commitment*, in MENTAL DISORDER AND CRIMINAL LAW 159, 168–69 (Robert F. Schopp et al. eds., 2009); Christina Lee, *Judicial Response to Psychopathic Criminals: Utilitarianism over Retribution*, 31 LAW & PSYCHOL. REV. 125, 127–32 (2007).

¹¹⁵ See HALL *supra* note 68, at 169 (“The use of ‘appreciate’ rather than ‘know’ conveys a broader sense of understanding than simple cognition.”).

¹¹⁶ See PARRY, *supra* note 44, at 311 (observing that “while the notion of appreciation of criminal conduct is distinguishable from the more narrow [sic] conception of knowing right

These claims find further confirmation in the second paragraph of the MPC test (the “caveat paragraph”), specifying that the notion of mental disease or defect does not include repeated manifestations of criminal or otherwise antisocial conduct.¹¹⁷ As previously noted, this provision was introduced with the explicit purpose of excluding psychopaths and, more broadly, patients suffering from socio-affective disorders from the range of eligible candidates for insanity defense.¹¹⁸ As Schopp observes:

Th[e] requirement [of emotional awareness] . . . would seem to exculpate the cold or vicious criminal who victimizes innocent people without experiencing sympathy or remorse. Yet, the insanity defense certainly is not intended to exculpate such criminals. Rather, these are just the people that the criminal law—and the prison system—are designed to deter.¹¹⁹

As can be noted, the rationale for excluding emotionally deficient perpetrators also lies in safeguarding public safety needs. Extreme forms of moral deviancy are considered symptomatic only of particularly dangerous personalities: the kinds that are very likely to harm society repeatedly.¹²⁰ There is consequently a tendency—in criminal law as well as in common understanding—to qualify patients suffering from pathological socio-affective deficits as iconic wrongdoers, as “evil” individuals who constantly and willingly reject and break the rules of societal coexistence.¹²¹ As such,

from wrong contemplated by the *M’Naughten* rule, the ALI model still emphasizes the offenders’ cognition of the criminality of their acts. In practice, the ALI model continues to rely upon instrumental reasoning capacities”).

¹¹⁷ MODEL PENAL CODE § 4.01(2) (AM. LAW INST. 1962) (“As used in this Article, the terms ‘mental disease or defect’ do not include an abnormality manifested only by repeated criminal or otherwise antisocial conduct.”)

¹¹⁸ See MODEL PENAL CODE § 4.01(2) (AM. LAW INST. 1962); see also SLOVENKO, *supra* note 62.

¹¹⁹ ROBERT SCHOPP, AUTOMATISM, INSANITY, AND THE PSYCHOLOGY OF CRIMINAL RESPONSIBILITY 33 (1991).

¹²⁰ See Edens, *supra* note 114.

¹²¹ See, e.g., Maria Isabel Gonzalez-Tapia et al., *A New Legal Treatment for Psychopaths? Perplexities for Legal Thinkers*, 54 INT. J. L. & PSYCHIATRY 46, 47 (2017) (“Psychopathy . . . represents archetypes of ‘evil’, of incorrigible criminals, for whom a retributive culpability-based punishment is not enough and a consequentialist ‘dangerousness-based’ legal response would be required.”); William Waller, “*Criminal Insanity & Public Morality*,” 4 WASH. U. JUR. REV. 183, 190–91 (2011) (observing that sociopaths [meaning people suffering from antisocial personality disorder] are not good candidates for the insanity defense, even if on the literal terms of a majority of permutations it ought to apply. Apart from self-selection, sociopaths can be handled routinely under the criminal law because their mental illness is, diagnostically speaking, behavioral Because the defense serves to control retribution . . . society has enough reservation with applying the defense outside of situations where the accused possesses certain overt, graphic, physiological characteristics. Faced with a disorder defined only in terms of a tendency

they fully deserve blame and punishment, as well as the kind of criminal incapacitation that curbs their antisocial tendencies while keeping society safe.

B. LACK OF SELF-CONTROL AND EMOTIONAL DISTURBANCE

As discussed above, criminal law is mostly insensitive to whether a defendant is in control at the time of the crime, unless this lack of control is the result of a mental disease affecting his or her cognitive faculties of knowledge and understanding. Conversely, the legal relevance of an incapacity for self-control that is linked to emotional impairments does not suffice for exculpation. While criminal law accepts that emotions may overturn reason and control, emotions do not do so to the point of providing grounds for insanity and blamelessness.

This view has its foundations in the folk-legal conception that emotions are not essential for moral rationality.¹²² Rather, emotions are mental states that can and should be defeated by rational powers.¹²³ As phrased by Finkel and Parrott, “in the Law’s folk psychology theory as well as in its normative expectations, there is the belief that control over one’s emotions is psychologically possible and normatively expected.”¹²⁴ Thus, although criminal law treats emotions as mental states that may override reason and self-control, if a defendant acts under strong emotional influence—however pathological in nature—this condition is not treated as the kind of moral rationality defect that might exculpate him or her.

While lack of self-control due to emotional impairments cannot ground a total excuse, there are a few cases in which criminal law allows for mitigation due to lack of self-control resulting from an emotional breakdown. Among them, the most emblematic are the mitigating conditions in common law’s “heat of passion”¹²⁵ and the MPC’s EED defenses.¹²⁶ Both mitigations constitute specific forms of diminished

toward antisocial behavior, however, and of which sufferers are sensationalized as evil, society will revert to its baseline of the criminal sanction . . .”).

¹²² See *supra* Section II.

¹²³ See *supra* Section II.

¹²⁴ See FINKEL & PARROT, *supra* note 90, at 137.

¹²⁵ See Reid Griffith Fontaine, *Adequate (Non)Provocation and Heat of Passion as Excuse Not Justification*, 43 U. MICH. J.L. REFORM 27, 29–30 (2009) (observing that under the Provocation/ Passion partial excuse, “the defendant must demonstrate that (a) he was adequately provoked, (b) as a direct result of said provocation, he became emotionally charged such that he lost self-control, (c) not enough time to ‘cool off’ passed between provocation and killing, and (d) he did not, in fact, cool off prior to killing his victim(s)”).

¹²⁶ See MODEL PENAL CODE § 210.3(1)(b) (AM. LAW. INST. 1962). The EED is a codified and expanded version of the Heat of Passion doctrine which—despite its greater

capacity¹²⁷ that bring substantially lower penalties. It is important to outline that such mitigation is only available for reducing murder to manslaughter and not for any other crime.¹²⁸ Thus, if a person commits any crime other than homicide in a state of emotional disturbance, he or she is not able to raise this kind of defense at the trial stage to obtain a reduction of penalty.¹²⁹

Both the EED and the heat of passion doctrines are based on the commonsense-based intuition that persons in extreme emotional conditions generally do not intend and cannot control very much of anything.¹³⁰ That is, when there is strong emotion there cannot be deliberation. Based on this intuition, the law at least grants mitigation by recognizing that people who kill while in a state of extremely heightened emotion—with or without prior

narrowness—still remains the most largely mitigating doctrine adopted by U.S. jurisdictions. Like the heat of passion standard, the EED standard allows that murder be reduced to voluntary manslaughter if the defendant, due to an extreme emotional breakdown, acted in an uncontrollable rage. Unlike heat of passion, however, the core component of the mitigation is that the killing must have been committed “under the influence of an extreme emotional disturbance for which there is a reasonable explanation and excuse,” regardless of whether the defendant was provoked or not. Therefore, “any affective experience sufficient to disable a person’s ‘usual intellectual controls’ or scrambles ‘normal rational thinking’ counts as an extreme emotional disturbance”. See *Patterson v. New York*, 432 U.S. 197, 206 (1977); *People v. Casassa*, 404 N.E. 2d 1310 (N.Y. 1980).

¹²⁷ See Paul Robinson, *Abnormal Mental State Mitigations of Murder – The U.S. Perspective*, in *LOSS OF CONTROL AND DIMINISHED RESPONSIBILITY: DOMESTIC, COMPARATIVE AND INTERNATIONAL PERSPECTIVE* 291 (Alan Reed & Michael Bohlander eds., 2011) (claiming that this mitigation is actually misleadingly referred to as diminished capacity or partial responsibility, in that it only negates the existence of an element of the crime but does not indicate reduced culpability or responsibility).

¹²⁸ Uri Moaz & Gideon Yaffe, *What Does Recent Neuroscience Tell Us About Criminal Responsibility?*, 3 *J. L. & BIOSCIENCE* 120, 136 (2016) (asserting that the reason why the law makes the EED mitigation available only for homicide lies in the fact that homicide is “a far less common crime than many others. This indicates the rather stingy attitude in the law towards basing differences in treatment on differences in control.”).

¹²⁹ Stephen J. Morse, *Diminished Rationality, Diminished Responsibility*, 1 *OHIO ST. J. CRIM L.* 289, 296 (2003) (“Why should these doctrines be limited to homicide? For example, suppose a defendant acting in the heat of passion intentionally burns the provoker’s property, rather than killing the provoker. Or suppose that an agent suffering from a non-culpable state of substantially diminished rationality commits arson. Some arsonists and some criminals generally might act with non-culpable, substantially impaired rationality that does not meet the standards for a full legal excuse. Compromised rationality and its effect on culpability are not limited to homicide. Fairness and proportionality require that doctrinal mitigation should be available in all cases in which culpability is substantially reduced.”).

¹³⁰ See Joshua Dressler, *Why Keep the Provocation Defense? Some Reflections on a Difficult Subject*, 86 *MINN. L. REV.* 959, 959 n.5 (2002) (“Provocation law is all about emotions.”).

provocation, depending on the standard—are less than fully in control of what they are doing.

An analysis of the doctrine of diminished capacity demonstrates an unequal treatment of lack of self-control due to emotional impairments, as compared to treatment of lack of control linked to cognitive defects.¹³¹ In line with this Article's thesis, the reason for this disparity is that cognition is considered to be the only essential source of moral rationality. Therefore, when cognitive faculties are impaired, a person can lose the capacity to control himself or herself and might be excused. Emotion, however, is viewed as the opposite of intellect and is thus not constitutive of the mental makeup of moral rationality. It follows that when there is no proof of a cognitive or intellectual defect, there is no space for a volitional excuse on the basis of emotional impairment alone. In the absence of a cognitive or intellectual defect, emotional impairment is simply not enough to satisfy the law.

III. EMOTIONS, MORALITY, AND ANTISOCIAL BEHAVIOR: INSIGHTS FROM NEUROSCIENCE

As illustrated previously, criminal law tends to deny emotions a positive role within the capacities for moral rationality that an individual must possess in order to be potentially criminally culpable. While the notion of culpability, and hence of culpability-related doctrines such as insanity, revolve around the sphere of cognition, emotions are mostly treated as sudden occurrences that can provoke a temporary distortion in an individual's capacity for practical moral reasoning.¹³² Unlike cognitive dysfunctions, emotional impairments are not treated as symptomatic of moral rationality defects.¹³³ Thus, emotional faculties receive no prominence in the substance of legal insanity.

The following sub-sections highlight that the legal overreliance on the sphere of cognition, as well as the marked disregard of emotions in the capacity for moral rationality, are empirically inaccurate. They do so by measuring the legal paradigm of the capacity for moral rationality against the neuroscientific (and behavioral) literature about the brain and the mental dynamics that underpin moral judgment and decision-making, and antisocial behavior. The aim here is not to conduct an exhaustive literature review of this body of neuroscientific studies. Rather, and more narrowly, it is to use relevant scientific literature to highlight the foremost mistaken

¹³¹ See discussion *supra* Section I.B.

¹³² See discussion *supra* Section II.

¹³³ *Id.*

legal-psychological assumptions emerging from the legal understanding of the capacity for moral rationality and, consequently, from insanity tests. First, “emotion” is as critical as “cognition” in moral decision-making and behavior. Second, the capacity for self-control does not necessarily depend on intellectual faculties of knowledge and understanding. Rather, self-control abilities depend on the balanced relationship between cognitive and emotional processes, both of which are critical in mediating and regulating our impulses and behavior. This dependence entails that a disruption in either the emotional or cognitive processes involved in self-control are equally capable of endangering one’s controlled choice of behavior. Thus, one can retain perfect factual and moral knowledge or understanding capacities, yet remain incapable of making adaptive choices or controlling one’s impulses.

A. “KNOWING” WITHOUT “FEELING” HAS A NEGATIVE IMPACT ON MORAL JUDGMENT AND DECISION-MAKING

The word “emotion” comes from the Latin *ex* (out) + *movere* (to move).¹³⁴ Its etymology is consistent with its core function, namely that emotions compel people to act and motivate them to act in a given way.¹³⁵ Psychologists and neuroscientists have developed numerous theories of emotions in order to understand the particular rationality of emotional reactions. These theories come in many varieties, but they largely impugn the view that emotions are mere primitive states that only distort reasoning.¹³⁶ Rather, they share the idea that emotions—both basic¹³⁷ and moral¹³⁸—provide critical guidance in reasoning and decision-making processes.

¹³⁴ *Emotion*, English Oxford Living Dictionaries, <https://en.oxforddictionaries.com/definition/emotion> [<https://perma.cc/GQ6G-BEYL>] (last visited June 16, 2018).

¹³⁵ Rene Rosfort & Giovanni Stanghellini, *How Do You Feel? Why Emotions Matter in Psychiatry*, 20 J. PSYCHOPATHOLOGY 381, 385 (2014) (“Emotions are the lived motivation for movement. Emotions are kinetic, dynamic forces that drive us in our ongoing interactions with the environment.”).

¹³⁶ See Jennifer Lerner et al., *Emotion and Decision Making*, 66 ANN. REV. PSYCHOL. 799 (2015); Elizabeth Phelps et al., *Emotion and Decision Making: Multiple Modulatory Neural Circuits*, 37 ANN. REV. NEUROSCIENCE 263 (2014).

¹³⁷ The category of basic emotions covers a disputed territory. By and large, basic emotions are discrete mental states including primitive emotions like happiness, sadness, fear, anger, and the like. See Paul Ekman, *Basic Emotions*, in HANDBOOK OF COGNITION AND EMOTION 45 (Tim Dalgleish & Mick J. Power eds., 1999).

¹³⁸ Moral emotions (also referred to as social emotions) differ from basic emotions, for they are intrinsically linked to the interests or welfare either of society as a whole or of persons other than the agent. See Jonathan Haidt, *The Moral Emotions*, in HANDBOOK OF AFFECTIVE SCIENCES 852 (Richard Davidson, Klaus Sherer, and H. Hill Goldsmith eds.,

According to most accredited scientific accounts, emotions serve a variety of functions within decision-making and behavior. First, and foremost, emotions have an *appraisal* (or evaluative) function:¹³⁹ that is, when people receive or perceive inputs (or stimuli, or events) from the environment, emotions act to provide meaning and value to the information being processed, and thus act as strong influences in pursuing appropriate behavior in response to that appraised information. Importantly, emotional appraisals strongly influence cognitive functions.^{140–141} Second, emotions are *motivational* states.¹⁴² Upon perception and appraisal of external stimuli, emotions help people select responses and thus motivate their behavior in responding to stimuli in an appropriate way.¹⁴³ Third, emotions are *adaptive*:¹⁴⁴ they help people prioritize and organize their behavior in ways that optimize their adjustment to the demands of the physical and social environment. Thus, emotions modulate people's behavioral responses by appropriately tuning their decisions to the demands or opportunities offered by the environment, thereby allowing them to

2003); see also Kathryn F. Jankowski & Hidehiko Takahashi, *Cognitive Neuroscience of Social Emotions and Implications for Psychopathology: Examining Embarrassment, Guilt, Envy, and Schadenfreude*, 68 PSYCHIATRY & CLINICAL NEUROSCIENCE 319 (2014) (“Social emotions are affective states elicited during social interactions and integral for promoting socially appropriate behaviors and discouraging socially inappropriate ones.”).

¹³⁹ See generally MAGDA ARNOLD, *EMOTION AND PERSONALITY* (1960); Richard S. Lazarus, *Thoughts on the Relation Between Emotion and Cognition*, 37 AM. PSYCHOL. 1019 (1982); Klaus R. Scherer, *Neuroscience Findings are Consistent with Appraisal Theories of Emotion; But Does The Brain “Respect” Constructionism?*, 35 BEHAV. & BRAIN SCI. 163 (2012).

¹⁴⁰ See Zhong Lin-Lu & Barbara Anne Doshier, *Cognitive Psychology*, 2 SCHOLARPEdia 2769 (2007) (explaining that in neuroscience, cognitive functions refer to cerebral activities that lead to knowledge, including all means and mechanisms of acquiring information. Cognitive functions encompass reasoning, memory, learning attention, and language and lead directly to the attainment of information and, thus, knowledge).

¹⁴¹ Mark D. Lewis & Rebecca M. Todd, *Getting Emotional: A Neural Perspective on Emotion, Intention, and Consciousness*, 12 J. CONSCIOUS STUD. 210, 212 (2005) (“[T]he biological function of emotion is to impel appropriate behavior, given past learning and precedent circumstances, by steering attention toward useful options for acting on the world and urging one to pursue them. Thus, cognition in general . . . is assumed to be guided by emotional relevance.”).

¹⁴² LAMBERT DECKERS, *MOTIVATION: BIOLOGICAL, PSYCHOLOGICAL, AND ENVIRONMENTAL* 367–98 (3d ed 2009); NICO H. FRIJDA, *Emotions and Action*, in FEELINGS AND EMOTIONS 158–73 (Antony S.R. Manstead, Nico Frijda & Agneta Fisher eds., 2004).

¹⁴³ FRIJDA, *supra* note 142 at 158–73; see also Carroll E. Izard, *Emotion Theory and Research: Highlights, Unanswered Questions, and Emerging Issues*, 60 ANN. REV. PSYCHOL. 1 (2009).

¹⁴⁴ Datcher Keltner & James J. Gross, *Functional Accounts of Emotions*, 13 COGNITION & EMOTION 467 (1999).

effectively apply the decision to take action.¹⁴⁵ Lastly, emotions are *regulatory*: they regulate the maintenance of internal bodily integrity, such that an organism can be prepared for specific reactions.¹⁴⁶ Thus, emotions regulate bodily responses by prompting specific reactions to external inputs.¹⁴⁷

Importantly, neuroscientific studies suggest that emotions play their guiding role in decision-making at both subconscious and conscious levels.¹⁴⁸ For a long time, an individual's emotional life was generally understood as a subconscious phenomenon belonging solely to the subcortical regions, the most primitive regions of the brain (notably, the limbic structures such as the amygdala),¹⁴⁹ while cognition was attributed to the highly developed, neocortical regions (notably, the prefrontal cortex [PFC]).¹⁵⁰ As such, scientific research long endorsed the view that the high functions of reasoning and decision-making were solely served by cognitive faculties.¹⁵¹

¹⁴⁵ *Id.* at 470 (“Emotions are adaptations to problems in the current human environment.”).

¹⁴⁶ See ANTONIO R. DAMASIO, *THE FEELING OF WHAT HAPPENS: BODY AND EMOTIONS IN THE MAKING OF CONSCIOUSNESS* 39 (1999) (asserting that emotions are part of the regulation of one's homeostasis, i.e. “the coordinated and largely automated physiological reactions required to maintain steady internal states in a living organism”). For example, when we are confronted with a threatening situation, the negative emotion of fear puts our brain and our body out of balance, for it provokes physiological changes such as rapid heartbeat and breathing. While provoking physiological changes, fear will predict the threat and prompt us to respond to and cope with that threat, and thus regain homeostasis by adjusting its physiological processes.

¹⁴⁷ *Id.*

¹⁴⁸ See *supra* note 8.

¹⁴⁹ See, e.g., James W. Papez, *A Proposed Mechanism of Emotion*, 79 ARCHIVES NEUROLOGY & PSYCHIATRY 725 (1937).

¹⁵⁰ See W. Gerrod Parrott & Jay Schulkin, *Neuropsychology and the Cognitive Nature of the Emotions*, 7 COGNITION & EMOTION 43 (1993).

¹⁵¹ See, e.g., Joseph E. LeDoux, *Emotions Circuits in the Brain*, 23 ANN. REV. NEUROSCIENCE 155, 156 (2000) (“Why did research on the brain mechanisms of emotion come to a halt after midcentury? . . . For one thing, emotion research was a victim of the cognitive revolution. The emergence of cognitive science shifted the interest of those concerned with the relation between psychological functions and neural mechanisms toward processes [perception and memory, for example] that were readily thought of in terms of computer-like operations Another factor that hindered work on emotions in neuroscience was that the problem of how the brain makes emotions seemed to have been solved in the early 1950s by the limbic system concept This appealing and convincing theory was the culmination of research on the brain mechanisms of emotion by many researchers, extending back to the late nineteenth century Studies of how the brain mediates cognitive processes seemingly had a long way to go to catch up with the deep understanding that had been achieved about emotions, and researchers flocked to the new and exciting topic of cognition and the brain to begin filling the gap.”).

In contrast to this outdated perspective, more recent studies have expanded the conception of the scope of the “emotional brain.”¹⁵² In particular, these studies have supported the view that emotional processes also involve significant participation of the neocortical regions—notably, the ventromedial prefrontal cortex (VmpFC) and the orbitofrontal cortex (OFC).¹⁵³ Studies on the neocortical dimension of emotional processes have led neuroscientists to also maintain a strong interrelation between cognitive and emotional processes in decision-making.¹⁵⁴ One crucial insight emerging from this line of research is that emotional and cognitive processes involved in decision-making often engage overlapping neural mechanisms.¹⁵⁵ Therefore, there is not a clear-cut distinction between “emotion” and “cognition” in decision-making processes in the brain, but “emotion” and “cognition” are strictly intertwined and they *equally* contribute to the production of decisions and behavior as a consequence.

Considering the significant involvement of emotion-related brain circuits in decision-making tasks, neuroscientists have come to attribute a fundamental role to emotions, when paired with cognitions, in helping to guide decision-making and behavioral outcomes. As Elizabeth Johnston and Leah Olson have suggested, “[t]he so called ‘cognitive’ brain functions, such as attention, perception, learning and memory, and decision-making can no longer be seen as separate and distinct from emotions; instead, they are inextricably infused with emotional assessments and feelings that accompany them.”¹⁵⁶ In a nutshell, the neuropsychological sciences have largely rejected the view that emotions are necessarily “disturbing factors”

¹⁵² See, e.g., Don Tucker et al., *Anatomy and Physiology of Human Emotion: Vertical Integration of Brainstem, Limbic, and Cortical Systems*, in HANDBOOK OF THE NEUROPSYCHOLOGY OF EMOTION 56 (Joan C. Borod ed., 2000).

¹⁵³ See *infra* 189–95.

¹⁵⁴ See, e.g., ANTONIO R. DAMASIO, DESCARTES’ ERROR: REASON, EMOTION, AND THE HUMAN BRAIN 54–79 (1994); Jennifer Lerner et al., *Emotion and Decision-Making*, 66 ANN. REV. PSYCHOL. 799, 802–11 (2015); Elizabeth A. Phelps et al., *Emotion and Decision Making: Multiple Modulatory Neural Circuits*, 37 ANN. REV. NEUROSCIENCE 263, 267–81 (2014); Megan Speer & Mauricio Delgado, *Emotion-Cognition Interactions in Memory and Decision-Making*, in STEVEN’S HANDBOOK OF DEVELOPMENTAL PSYCHOLOGY AND COGNITIVE NEUROSCIENCE 591, 596–605 (John Wixted & Eric Jan Wagermakers eds, 2018).

¹⁵⁵ See Richard J. Davidson, *Cognitive Neuroscience Needs Affective Neuroscience (and Vice Versa)*, 42 BRAIN & COGNITION 89, 91 (2000) (“Cognition would be rudderless without the accompaniment of emotion, just as emotion would be primitive without the participation of cognition.”).

¹⁵⁶ ELIZABETH JOHNSTON & LEAH OLSON, THE FEELING BRAIN: THE BIOLOGY AND PSYCHOLOGY AND EMOTIONS 307 (2015).

for decision-making or that emotions can possibly be factored out of the decision-making processes.¹⁵⁷

The influential role of emotions in guiding deliberations and behavior, as well as the strict interrelation between emotional and cognitive functions in decision-making processes, have been further contextualized in moral judgments. A growing body of behavioral literature has consistently indicated that emotions—most notably, certain moral emotions¹⁵⁸—are the core driving force of the deliberative processes involved in moral judgments.¹⁵⁹ By and large, moral emotions help people recognize and appreciate the moral value of given morally salient stimuli (e.g., good or bad, right or wrong).¹⁶⁰ Based on this evaluation, moral emotions motivate and orient people’s reactions to those stimuli and consequently help them adapt decisional and behavioral responses.¹⁶¹ Ultimately, moral emotions profoundly influence one’s “adherence (or lack of adherence) to moral standards.”¹⁶²

Different moral emotions convey different information about various perceived events and shape moral judgments by prioritizing different socio-moral concerns.¹⁶³ Hence, moral emotions affect moral decisions and behavior in varying ways.¹⁶⁴ For example, some studies have shown the

¹⁵⁷ *Id.*

¹⁵⁸ Notwithstanding the various classifications that have been proposed, moral emotions are by and large classified in four main categories: self-conscious moral emotions (including guilt, shame, embarrassment, and pride); other-condemning emotions (anger, disgust, contempt); other-praising emotions (love, elevation); other-suffering emotions (compassion). See Haidt, *supra* note 138.

¹⁵⁹ See *infra* notes 165–63.

¹⁶⁰ See Dacher Keltner et al., *Emotions as Moral Intuitions*, in *AFFECT IN SOCIAL THINKING AND BEHAVIOR* 161, 164–68 (Joseph P. Forgas ed., 2006).

¹⁶¹ See Rimma Teper et al., *How Emotions Shape Moral Behavior: Some Answers (and Questions) for the Field of Moral Psychology*, 91 *SOC. & PERSONALITY PSYCHOL. COMPASS* 1, 5 (2015) (“Within the past decade, psychologists have theorized about the ways in which these emotions might drive moral decision making, in both their real and anticipated forms For instance, people might be motivated to relieve the pre-decisional negative affect (e.g. guilt), avoid post-decisional anticipated negative affect (e.g. shame), or achieve post-decisional positive affect In other words, moral emotions can provide both the information and motivational force to do the ‘right thing’”).

¹⁶² See June Price Tangney et al., *Moral Emotions and Moral Behavior*, 58 *ANN. REV. PSYCHOL.* 345, 347 (2007).

¹⁶³ See, e.g., Elizabeth J. Horberg et al., *Emotions as Moral Amplifiers: An Appraisal Tendency Approach to the Influences of Distinct Emotions upon Moral Judgment*, 3 *EMOTION REV.* 237, 238 (2011).

¹⁶⁴ *Id.*

crucial relevance of guilt,¹⁶⁵ in combination with other-oriented empathy, in inhibiting antisocial impulses and promoting prosocial behavior.¹⁶⁶ Other work has connected compassion to concerns about caring for and reducing harm to others, particularly those in need.¹⁶⁷ Compassion is aroused by perceptions of the need, suffering, or weakness of others, and motivates prosocial action even if it is costly to the self.¹⁶⁸ Research has found that compassion prominently shapes moral judgments of harm and care.¹⁶⁹

Other studies have instead focused on the relationship between moral judgments and empathy.¹⁷⁰ Notwithstanding the lack of a unanimous definition of empathy, it can be defined as a multi-dimensional¹⁷¹ socio-affective process consisting of “gaining information about the internal affective representations of others,”¹⁷² thereby eliciting vicarious emotional responses.¹⁷³ Although the way empathy—and, notably, each of its components, namely emotional sharing, perspective-taking, and empathic concern¹⁷⁴—interacts with and influences morality is debated,¹⁷⁵ convincing

¹⁶⁵ See June P. Tangney et al., *Shame, Guilt and Remorse: Implications for Offender Populations*, 22 J. FORENS. PSYCHIATRY & PSYCHOL. 706 (2011).

¹⁶⁶ See Linda Torstveit et al., *Empathy, Guilt Proneness and Gender: Relative Contributions to Prosocial Behaviour*, 12 EUR. J. PSYCHOL. 260, 265–66 (2016); Jeffrey Stewing et al., *Shaming, Blaming, and Maiming: Functional Links Among the Moral Emotions, Externalization of Blame, and Aggression*, 44 J. RES. PERSONALITY 91 (2010).

¹⁶⁷ See, e.g., Helen Y. Weng et al., *The Role of Compassion in Altruistic Helping and Punishment Behavior*, 10 PLOS ONE 1, 12 (2015).

¹⁶⁸ Jennifer L. Goetz et al., *Compassion: An Evolutionary Analysis and Empirical Review*, 136 PSYCHOL. BULL. 351, 356 (2010).

¹⁶⁹ See *id.* at 354.

¹⁷⁰ See generally *infra* notes 177–76.

¹⁷¹ See Mark H. Davis, *Measuring Individual Differences in Empathy: Evidence for a Multidimensional Approach*, 44 J. PERSONALITY & SOC. PSYCHOL. 113 (1983).

¹⁷² See Giuseppe Ugazio et al., *Are Empathy and Morality Linked? Insights from Moral Psychology, Social and Decision Neuroscience, and Philosophy*, in *EMPATHY AND MORALITY* 155, 161 (Heidi Maibom ed., 2014).

¹⁷³ See Nancy Einseberg et al., *Empathy-Related Responding: Associations with Prosocial Behavior, Aggression, and Intergroup Relations*, 4 SOC. ISSUES & POL’Y REV. 143 (2010).

¹⁷⁴ Emotional sharing (sometimes referred to as empathic arousal or emotional contagion) reflects the natural capacity to become affectively aroused by others’ emotions. Perspective-taking is the ability to consciously put oneself into the mind of another individual and imagine what that person is thinking or feeling. Empathic concern corresponds to the motivation of caring for another’s welfare. Critically, neuroscientific research suggests that each of these facets of empathy emerges from specific neurobiological processes—both emotional and cognitive. Thus, it is very likely that each empathy facet uniquely influences moral cognition and predicts differential outcomes in moral behavior. See, e.g., Raeanne C. Moore et al., *Distinct Neural Correlates of Emotional and Cognitive Empathy in Older Adults*, 232 PSYCHIATRY RES.: NEUROIMAGING 42 (2015); Sylvia Morelli

evidence exists supporting the claim that empathy does have an overall bearing on moral judgments.¹⁷⁶ Especially concerning other individuals, empathy promotes caregiving behavior, and it also leads to aversions to violence.¹⁷⁷

Some authors have suggested a particularly key role for “empathic concern”—that is, the motivation of caring for another’s welfare—in informing moral judgments, for empathic concern is largely involved in altruistic behavior in response to someone in distress.¹⁷⁸ Studies using the “trolley dilemma,”¹⁷⁹ an iconic experiment in moral philosophy, have indicated that people with low levels of empathic concern are more likely to endorse utilitarian solutions to personal moral dilemmas: solutions that require personally harming someone in order to achieve the greater good.¹⁸⁰ Importantly, neuroscientific studies testing this hypothesis have related empathic concern to activity in the amygdala and the VmPFC that, as

et al., *The Neural Basis of Empathy for Components of Empathy: Predicting Daily Prosocial Behavior*, 9 SOC. COGNITIVE & AFFECTIVE NEUROSCIENCE 39 (2014); Jamil Zaki & Kevin N. Ochsner, *The Neuroscience of Empathy: Progress, Pitfalls, and Promise*, 15 NATURE NEUROSCIENCE 675 (2012).

¹⁷⁵ See, e.g., Jean Decety & Jason M. Cowell, *Friends or Foes: Is Empathy Necessary for Moral Behavior?*, 9 PERSP. ON PSYCHOL. SCI. 525, 530 (2015).

¹⁷⁶ See, e.g., Jamil Zaki, *Empathy is a Moral Force*, in ATLAS OF MORAL PSYCHOLOGY 49 (Kurt Gray & Jesse Graham eds., 2018); Geert-Jan Will & Eduard T. Klapwijk, *Neural Systems Involved in Moral Judgment and Moral Action*, 34 J. NEUROSCIENCE 10459 (2014); Keith Yoder & Jean Decety, *The Neuroscience of Morality and Social Decision-Making*, 24 PSYCHOL., CRIME & L. 279 (2018); Fang Cui et al., *Moral Judgment Modulates Neural Responses to the Perception of Other’s Pain: An ERP Study*, 6 SCI. REP. 1 (2016).

¹⁷⁷ See Jean Decety & Jason M. Cowell, *Empathy, Justice, and Moral Behavior*, 6 AJOB NEUROSCIENCE 3 (2015); see also Ugazio et al., *supra note 172*, at 165 (noting that “[e]mpathy, by eliciting feelings of approbation or disapprobation, can be used to decide whether an action should be considered morally right or wrong Furthermore, by making a person aware of the emotional state of others, empathy can motivate people to judge and eventually act accordingly. For instance, if someone is in a negative emotional state as a result of another person’s actions, for example, feels pain after being hit by another person, empathy may motivate an observer to judge that hitting others is morally wrong and, by extension, may motivate him to help the victim.”).

¹⁷⁸ See, e.g., Oriol FeldmanHall et al., *Empathic Concern Drives Costly Altruism*, 105 NEUROIMAGE 347, 352–54 (2015).

¹⁷⁹ Judith Jarvis Thomson, *Killing, Letting Die, and the Trolley Problem*, 59 THE MONIST 204 (1976).

¹⁸⁰ Ezequiel Gleichgerrcht & Liane Young, *Low Levels of Empathic Concern Predict Utilitarian Moral Judgment*, 8 PLOS ONE e60418 (2013); Indrajeet Patil & Giorgia Silani, *Reduced Empathic Concern Leads to Utilitarian Moral Judgments in Trait Alexithymia*, 5 FRONT. PSYCHOL. 501 (2014); Liane Young et al., *Damage to Ventromedial Prefrontal Cortex Impairs Judgment of Harmful Intent*, 65 NEURON 845 (2010).

discussed in more detail shortly, are key brain regions of the socio-emotional circuitry involved in moral judgments.

The insights into the relationship among moral emotions, empathy, and moral judgments are further complemented by neuroscientific advances regarding the neural correlates¹⁸¹ of moral judgment and decision-making. While brain scans alone cannot exactly determine which emotions are at stake in moral judgments and decision-making, studies using neuroimaging techniques “have consistently implicated those brain regions implicated in emotional processing, including moral emotion processing.”^{182–183} Importantly, the growing body of these studies has led several neuroscientists to indicate the existence of a “neuromoral network,”¹⁸⁴ a network of brain areas that appear to be constantly and significantly involved in moral judgment and decision-making. According to the prevailing view, the best-replicated neural correlates of morality broadly recruit a *fronto-temporo-subcortical* network, and therefore comprise both cognitive and emotional components.¹⁸⁵ The “cognitive” components of the moral circuit are mainly localized in the dorsolateral prefrontal cortex

¹⁸¹ Neural correlates can be defined as the neuronal mechanisms that correspond with—i.e., correlate directly with—a particular experience. For instance, research refers to the neural correlates of consciousness as “the minimum neural mechanisms sufficient for any one specific conscious percept.” See Christof Koch et al., *Neural Correlates of Consciousness: Progress and Problems*, 17 NAT. REV. NEUROSCI. 307 (2016).

¹⁸² James Blair & Katherine Fowler, *Moral Emotions and Moral Reasoning from the Perspective of Affective Cognitive Neuroscience: A Selective Review*, 2 EUR. J. DEV. SCI. 303, 314 (2008).

¹⁸³ Jorge Moll et al., *Frontopolar and Anterior Temporal Cortex Activation in a Moral Judgment Task: Preliminary Functional MRI Results in Normal Subjects*, 59 ARQ NEUROPSIQUIATR 657 (2001); Jorge Moll et al., *Functional Networks in Emotional Moral and Nonmoral Social Judgments*, 16 NEUROIMAGE 696 (2002); Joshua Greene et al., *An fMRI Investigation of Emotional Engagement in Moral Judgment*, 293 SCIENCE 2105 (2001); Joshua Greene et al., *The Neural Bases of Cognitive Conflict and Control in Moral Judgment*, 44 NEURON 389 (2004).

¹⁸⁴ See Mario F. Mendez, *The Neurobiology of Moral Behavior: Review and Neuropsychiatric Implications*, 14 CNS SPECTR. 608 (2009); Leo Pascual et al., *How Does Morality Work in the Brain? A Functional and Structural Perspective of Moral Behavior*, 7 FRONT. INTEGR. NEUROSCIENCE 1 (2013); Kristine Prehn & Hauke Heekeren, *Moral Judgment and the Brain: A Functional Approach to the Question of Emotion and Cognition in Moral Judgment Integrating Psychology, Neuroscience and Evolutionary Biology*, in THE MORAL BRAIN: ESSAYS ON THE EVOLUTIONARY AND NEUROSCIENTIFIC ASPECTS OF MORALITY 129 (Johan Braeckman, Jan Verplaetse & Jelle De Schrijver eds., 2009); Joshua Greene & Jonathan Haidt, *How (and Where) Does Moral Judgment Work?*, 6 TRENDS COGN. SCI. 517 (2002).

¹⁸⁵ See Mendez, *supra* note 184, at 608–09.

(DLPFC)¹⁸⁶ and the anterior cingulate cortex (ACC),¹⁸⁷ the “emotional” components of the moral circuitry instead involve the amygdala,¹⁸⁸ the OFC,¹⁸⁹ and notably the VmPFC,¹⁹⁰ which is also thought of as the “integrative center for innate morality.”¹⁹¹ The latter areas support specific affective aspects of moral judgment and decision-making, including emotional perception, sensitivity to reward and punishment, and motivation, especially in so-called “care-based morality.”¹⁹²

Roughly, studies suggest that the socio-emotional mechanism involved in moral judgment and decision-making goes as follows: perceived or received morally salient stimuli are first processed at the sub-cortical level by certain structures of the limbic system, in particular the amygdala.¹⁹³ Upon receiving a stimulus, the amygdala attaches either a positive or a negative emotional valence to it (e.g., good or bad), which is represented as an outcome valence within the VmPFC and the OFC.¹⁹⁴ The VmPFC “mediates automatic moral and ‘prosocial’ reactions, such as discomfort at the prospect of being a direct agent of a personal moral violation or of harm to someone else,”¹⁹⁵ thereby becoming very active in moral emotions that are positively linked to prosocial behavior, such as guilt or compassion.

¹⁸⁶ See Kristine Prehn et al., *Individual Differences in Moral Judgment Competence Influence Neural Correlates of Socio-Normative Judgments*, 3 SOC. COGNITIVE & AFFECTIVE NEUROSCIENCE 33 (2008).

¹⁸⁷ *Id.*

¹⁸⁸ See James Blair, *The Amygdala and the Ventromedial Prefrontal Cortex in Morality and Psychopathy*, 11 TRENDS COGN. SCI. 387 (2007).

¹⁸⁹ See Pascual, *supra* note 184.

¹⁹⁰ See Blair, *supra* note 188; Liane Young & Michael Koenigs, *Investigating Emotion in Moral Cognition: A Review of Evidence from Functional Neuroimaging and Neuropsychology*, 84 BRIT. MED. BULL. 69 (2007); Chuanpeng Hu & Xiaoming Jiang, *An Emotion Regulation Role of Ventromedial Prefrontal Cortex in Moral Judgment*, 8 FRONT. HUM. NEUROSCI. 873 (2014).

¹⁹¹ Donatella Marazziti et al., *The Neurobiology of Moral Sense: Facts or Hypotheses?*, 12 ANNALS GEN. PSYCHIATRY 6 (2013).

¹⁹² See Blair, *supra* note 188 (describing care-based morality as “those forms of moral reasoning that concern actions that harm others”).

¹⁹³ See, e.g., Pascual, *supra* note 184; Amitai Shenhav & Joshua Greene, *Integrating Moral Judgment: Dissociating the Roles of the Amygdala and Ventromedial Prefrontal Cortex*, 34 J. NEUROSCI. 4741 (2014); Manuela Fumagalli & Alberto Priori, *Functional and Clinical Neuroanatomy of Morality*, 135 BRAIN (2006).

¹⁹⁴ See, e.g., Pascual, *id.*; Liane Young & Michael Koenigs, *Investigating Emotion in Moral Cognition: A Review of Evidence from Functional Neuroimaging and Neuropsychology*, 84 BR. MED. BULL. 69 (2007); Abigail A. Marsh et al., *Reduced Amygdala-Orbitofrontal Connectivity during Moral Judgments in Youths with Disruptive Behavior Disorders and Psychopathic Traits*, 194 PSYCHIATRY RES. 279 (2011).

¹⁹⁵ Mendez, *supra* note 184, at 610.

Also, the VmPFC attributes moral and emotional values to social stimuli, anticipates their future outcome, and modulates the mechanisms of empathy¹⁹⁶ and perception of others' intentions. While the VmPFC serves mostly an *evaluative* function in the processing of emotional stimuli, the OFC serves the function of *filtering* emotional stimuli, dampening arousal to irrelevant inputs, and maintaining neural focus on task-relevant associations.¹⁹⁷ In so doing, the OFC mediates aversive responses related to the social context, modifies responses based on feedback, and inhibits automatic behavior triggered by the amygdala.

The critical role of socio-emotional brain circuits in moral judgment and decision-making has received further support in lesion studies. For instance, Sobhani and Bechara¹⁹⁸ indicated that anatomical lesions to and dysfunctions of the VmPFC and its reciprocal connections with the amygdala appear to lead to a lack of empathic concern, a diminished sense of guilt and inappropriate social behavior. In a study on VmPFC patients, Koenigs et al.¹⁹⁹ administered personal and non-personal moral dilemmas tests to VmPFC-damaged patients. While these patients provided the same utilitarian responses to non-personal harm dilemmas as the control subjects, they showed a far more marked utilitarian reasoning in personal moral dilemmas, compared to the control subjects.²⁰⁰ In addition to confirming the crucial role of the VmPFC in attaching emotional valences to moral considerations, these researchers also observed that normal utilitarian reasoning in impersonal dilemmas confirms that VmPFC patients retain intact cognitive intellectual abilities.²⁰¹ Likewise, Mario Mendez et al.²⁰² confirmed the crucial role of the VmPFC in moral judgment and decision-making. They found that patients with frontotemporal dementia, who are characterized by a lack of empathic concern and likelihood to engage in

¹⁹⁶ See *supra* note 181.

¹⁹⁷ Joshua Knabb et al., *Neuroscience, Moral Reasoning, and the Law*, 27 BEHAV. SCI. L. 219, 221 (2009).

¹⁹⁸ Mona Sobhani & Antoine Bechara, *A Somatic Marker Perspective of Immoral and Corrupt Behavior*, 6 SOC. NEUROSCI. 640, 640–41 (2011).

¹⁹⁹ Michael Koenigs et al., *Damage to the Prefrontal Cortex Increases Utilitarian Moral Judgments*, 446 NATURE 908 (2007).

²⁰⁰ *Id.* at 910 (“VMPC patients’ judgements differed from comparison subjects’ only for the high-conflict personal moral dilemmas, all of which featured competing considerations of aggregate welfare on the one hand, and, on the other hand, harm to others that would normally evoke a strong social emotion. Low-conflict personal moral scenarios lacked this degree of competition.”).

²⁰¹ *Id.* (“[T]he current results suggest that the VMPC is a critical neural substrate for the intuitive/affective but not for the conscious/rational system.”).

²⁰² Mario Mendez et al., *An Investigation of Moral Judgement in Frontotemporal Dementia*, 18 COGN. BEHAV. NEUROL. 193 (2005).

antisocial conduct, also tended to favor the utilitarian action in personal moral dilemmas.²⁰³

Additional support for this view has been provided by neuroimaging studies conducted on psychopathic individuals.²⁰⁴ Raine and Yang²⁰⁵ reviewed and integrated literature on moral reasoning and antisocial behavior to trace a neural moral model of antisocial behavior with a focus on psychopathy. Interestingly, their study found that the key brain areas that are either structurally or functionally impaired in patients suffering from psychopathy (e.g., the VmPFC or the amygdala) correspond to those being part of the socio-emotional circuitry of the neuromoral network (i.e., areas that are significantly involved in emotional processing, including moral emotion processing).²⁰⁶ Impairments in these brain areas seem to explain why patients affected by this type of disorder show severe empathy deficits (notably, a lack of empathic concern), callousness or emotional flatness, and deficiencies in using emotional information to regulate their behavior and to respond to other individuals' distress.²⁰⁷ Importantly, individuals with psychopathy do seem to retain a "cognitive" understanding of moral and legal wrongs.²⁰⁸ However, as one study indicated, psychopaths treat the word "wrong" in a purely conventional way, as if it

²⁰³ *Id.* at 195–96 (“The FTD patients differed in the way they responded to an emotionally based ‘personal’ moral dilemma compared with the AD patients and normal control subjects. The FTD patients retained knowledge of moral rules and norms and could reason about the right and wrong of a situation. In contrast, they appeared to have diminished emotional identification with others and solved moral dilemmas in an impersonal fashion.”).

²⁰⁴ The neuroscientific literature on psychopathy is very wide. *See generally* KENT A. KIEHL & WALTER SINNOTT-ARMSTRONG, *HANDBOOK ON PSYCHOPATHY AND LAW* (2013); Kent A. Kiehl et al., *Limbic Abnormalities in Affective Processing by Criminal Psychopaths as Revealed by Functional Magnetic Resonance Imaging*, 50 *BIOL. PSYCHIATRY* 677 (2001); R.J.R. Blair, *Responding to the Emotions of Others: Dissociating Forms of Empathy Through the Study of Typical and Psychiatric Population*, 14 *CONSCIOUSNESS & COGN.* 698 (2005); Yaling Yang et al., *Morphological Alterations in the Prefrontal Cortex and the Amygdala in Unsuccessful Psychopaths*, 119 *J. ABNORM. PSYCHOL.* 546 (2010).

²⁰⁵ Adrian Raine & Yaling Yang, *Neural Foundations to Moral Reasoning and Antisocial Behavior*, 1 *SOC. COGN. & AFFECTIVE NEUROSCI.* 203 (2006).

²⁰⁶ *Id.* at 205–06.

²⁰⁷ *See, e.g.*, Jean Decety et al., *Brain Response to Empathy-Eliciting Scenarios Involving Pain in Incarcerated Individuals with Psychopathy*, 70 *JAMA PSYCHIATRY* 638 (2013).

²⁰⁸ Neil Levy, *The Responsibility of the Psychopath Revisited*, 14 *PHIL. PSYCHIATRY & PSYCHOL.* 129, 132 (2007) (claiming that psychopaths know their actions “are widely perceived to be wrong . . . they are unable to grasp the distinctive nature and significance of their wrongness”).

simply meant prohibited by local authorities.²⁰⁹ This means that even if psychopaths are perfectly capable of *propositionally* distinguishing rightness from wrongness, and therefore of knowing the moral meaning of a given action, they tend to make socially poor or immoral decisions because they are unable to generate the feelings that guide adaptive decision-making in healthy individuals.

All things considered, neuroscientific studies seem to confirm the substantial artificiality of the net distinction between “cognition” and “emotion” in moral judgment and decision-making. First, emotion is not essentially disruptive to moral reasoning; rather, it is part of it. Second, knowledge or understanding of moral wrongs does not lead to moral behavior if it exists *without* emotional influence. That is, appreciating the moral meaning of one’s own conduct does not simply require verbal knowledge of what is right and what is wrong, but also a *feeling* of the moral significance of given conduct in a given social context.²¹⁰

Changing the perspective on moral (and antisocial) decision-making may have significant implications for criminal law’s rationalist conception of the culpable agent. If emotions and feelings are integral parts of the mental processes leading to moral decisions and behaviors, a truly accurate model of the legally relevant mind that underpins culpability-related doctrines such as insanity should also include the emotional dimension in its relevant substance. Emotion and affect have a bearing on moral rationality, and the cognitive ability to know the value of an action does not sufficiently explain one’s capacity to behave morally in a given social context without also considering emotional influences. Considering these insights, a rational moral agent is one who is also able to *emotionally appreciate*—rather than simply *know*—the moral significance of his or her conduct.²¹¹ When this capacity is severely impaired, the agent falls short of being a morally rational individual. Ideally, he or she should be excused.

²⁰⁹ R.J.R. Blair, *A Cognitive Developmental Approach to Morality: Investigating the Psychopath*, 57 *COGNITION* 1 (1995).

²¹⁰ See Raine & Yang, *supra* note 205, at 209 (“[I]t is predominantly the *feeling* of what is moral that is deficient in antisocial groups, rather than the *knowing* of what is moral. This moral feeling, centered on the PFC and amygdala, is the engine that translates the cognitive recognition that an act is immoral into behavioral inhibition—and it is this engine that functions less well in antisocial . . . individuals.”).

²¹¹ See Ralph Slovenko, *Psychiatry, Criminal Law, and the Role of the Psychiatrist*, 12 *DUKE L.J.* 395, 397 (“The dictum, ‘Cogito, ergo sum,’ ‘I think, therefore I am,’ is, however, the formula for the schizoid intellectual’s struggle to possess an ego. A healthy human being would be more likely to start from ‘I feel, therefore I am’.”).

B. SELF-CONTROL ABILITIES ALSO DEPEND ON EMOTIONS

Neuroscientific studies also dispute the legal conception of the (lack of) capacity for self-control. Neuroscientific research on self-control is ongoing, and a clear-cut taxonomy of self-control is lacking,²¹² but there is a consensus understanding of self-control rather as a multidimensional construct²¹³ involving a number of distinct cognitive and socio-emotional capacities, each of which contributes to individuals' control abilities.²¹⁴

Regarding moral contexts, neuroscientific research suggests that people's capacity to regulate and inhibit impulses also—though not exclusively—depends on their ability to “wilfully suspend”²¹⁵ immediate gratifications in favor of long-term outcomes, which is critical to pro-social, law-abiding behavior.²¹⁶ This ability appears to depend on the *reward*

²¹² Joshua Buckholtz et al., *A Neuro-Legal Lingua Franca: Bridging Law and Neuroscience on the Issue of Self-Control*, 5 MENTAL HEALTH L. & POL'Y. J. 1, 13 (2016). Admittedly, behavioral and neuroscientific terminology regarding self-control manifests a lack of conceptual clarity. Looking at the literature, the notion of self-control is either related to or distinguished from other abilities including cognitive control, self-regulation, or emotion regulation. Each of these abilities appears to have a role in buffering impulsive behavior. *See infra* notes 214–16. In this Section, I mostly refer to what psychological and neuroscientific literature frequently addresses as “self-regulation,” and to its role in impulse and behavioral control.

²¹³ *Id.* at 13.

²¹⁴ *See* Marc Lewis & Rebecca Todd, *The Self-Regulating Brain: Cortical-Subcortical Feedback and the Development of Intelligent Action*, 22 COGN. DEV. 406 (2007); Ethan Kross & Kevin Ochsner, *Integrating Research on Self-control Across Multiple Levels of Analysis: Insights from Social Cognitive and Affective Neuroscience*, in SELF CONTROL IN SOCIETY, MIND, AND BRAIN 76 (Ran Hassin, Kevin Ochsner, & Yaacov Trope eds., 2010); Sezin Öner, *Neural Substrates of Cognitive Emotion Regulation: A Brief Review*, 28 PSYCHIATRY & CLIN. PSYCHOPHARMACOLOGY 91 (2017).

²¹⁵ *See* Cole Korponay & Michael Koenigs, *The Neurobiology of Antisocial and Amoral Behavior*, in LEGAL INSANITY AND THE BRAIN 27 (Sofia Moratti & Dennis Patterson eds., 2016).

²¹⁶ *Id.* In psychological and neuroscientific literature, this ability is frequently referred to as “self-regulation”: *see, e.g.*, Lambros Lazarus et al., *The Roles of Impulsivity, Self-regulation, and Emotion Regulation in the Experience of Self-disgust*, 43 MOTIV. & EMO. 145, 147 (2019) (defining self-regulation as “people’s capacity to focus on their long-term goals and resist temptation and impulses for immediate gratification . . . As such, self-regulation involves the ability to alter thoughts, actions, and emotions in a way that serves goal striving, whether the goal is set by the self, the society or both.”). While some authors use the term self-regulation and self-control interchangeably, others consider self-control to be a subset of self-regulation or vice versa. *See, e.g.*, Wilhelm Hoffmann et al., *Executive Functions and Self-regulation*, 16 TRENDS IN COGN. SCI., 174, 174 (2012) (defining self-regulation as “goal-directed behavior” and self-control as “a narrower subset of self-regulatory processes [aiming] to override unwanted, prepotent impulses or urges . . .”). *Cfr.* Buckholtz et al. *supra* note 212 at 15–16 (referring to the

system, which involves and connects the activity of both prototypical emotional subcortical (including the amygdala, the ventral tegmental area and the *nucleus accumbens*) and neocortical brain areas (notably the DLPFC, dorsal-ACC [dACC], the OFC, and the VmPFC), both of which serve interdependent functions of emotional processing, motivation, and inhibition in goal-directed behavior.²¹⁷ Studies indicate that subcortical regions involved in the reward system, such as the amygdala, process perceptual stimuli (events) through the generation of appropriate emotional responses.²¹⁸ Consequently, the prefrontal regions exert modulatory control on the representation of perceptual stimuli, and thus in turn modulate the expression and evaluation of emotions.²¹⁹ Put this way, the emotional response to a given event (or stimulus) facilitates the exertion of cognitive control. In its turn, cognitive control—also referred to as executive functions—appears to downregulate emotional inputs.

As emerges from this brief explanation, there is nothing in the brain resembling an absolute regulator and an absolute regulatee of impulses and behavior. Rather, impulse and behavioral regulation are governed by a sufficiently powerful coordination system between subcortical and cortical brain structures and, thus, they largely involve cognitive and emotional functions.²²⁰ As Lewis and Todd suggested, cortical and subcortical regions stand in a reciprocal relationship during decision-making and self-regulation in that “cortical activities regulate subcortical activities through executive modulation of prepotent appraisals and emotional responses; [while] subcortical systems regulate the cortex by tuning its activities to the demands or opportunities provided by the environment.”²²¹ Therefore, emotional appraisal of external stimuli, and cognitive control over emotional responses are both critical in guaranteeing appropriate choices and thus behavioral outcomes.

ability to delay immediate gratification in favor of long-term outcomes as a specific domain of self-control).

²¹⁷ *Id.*

²¹⁸ See, e.g., Elizabeth Murray, *The Amygdala, Reward and Emotion*, 11 TRENDS COGN. SCI. 489 (2007).

²¹⁹ See, e.g., Öner, *supra* note 214; Liyang Sai et al., *Individual Differences in the Habitual Use of Cognitive Reappraisal Predict the Reward-Related Processing*, 6 FRONT. PSYCHOL. 1256 (2015); J.T. Buhle et al., *Cognitive Reappraisal of Emotion: A Meta-Analysis of Human Neuroimaging Studies*, 24 CEREB. CORTEX 2981 (2014).

²²⁰ See Lewis & Todd, *supra* note 214, at 412 (“Coordination across levels of the neuroaxis is a powerful vehicle for self-regulation, and specifying particular structures at particular levels opens the door to a precise descriptive language for modeling psychological self-regulation using neural terms.”).

²²¹ *Id.* at 406.

Importantly, the systems involved in inhibition and impulse regulation may be irrespective of an individual's moral knowledge or understanding of his or her conduct. As Sapolsky described it, "it is possible for a person to retain the cognitive capacity to distinguish right from wrong behaviour and, nonetheless, for reasons of mental illness, to be organically incapable of regulating the appropriateness of their behaviour."²²² Confirming this account, lesion studies have suggested that people with abnormalities in the functioning of the emotion-related brain regions involved in the reward system—like the VmPFC—possess adequate social and moral knowledge, but appear to be unable to effectively apply that knowledge to action.²²³ Consequently, even if they are able to state what they should do in a given moral situation, they choose to do something else.

As Miller et al. observed, these findings suggest that the VmPFC functions as the "site of interaction" of "valuation and self-control processes"²²⁴ during decision-making. Such interaction seems "to facilitate successful self-control."²²⁵ Importantly, these results fit into the larger body of data indicating the important role of the VmPFC in mediating and integrating cognitive and emotional influences on decision-making and behavior.²²⁶ Antonio Damasio's famous work on his young patient, to whom he referred as Elliot, supports the latter claim. Elliot had undergone a radical personality change after a surgery to remove a brain tumor on the surface of his frontal lobes.²²⁷ Elliot's intelligence had remained substantially intact after the operation.²²⁸

²²² Robert Sapolsky, *The Frontal Cortex and the Criminal Justice System*, 359 PHIL. TRANS. R. SOC'Y. LOND. B. 1787, 1790 (2004).

²²³ See, e.g., Young et al., *supra* note 180; Rupa Gupta et al., *The Amygdala and Decision-Making*, 49 NEUROPSYCHOLOGIA 760 (2011).

²²⁴ Erica Miller et al., *Delay Discounting: A Two-Systems Perspective*, HANDBOOK OF EMOTION REGULATION 93, 102 (James Gross ed., 2d ed 2014).

²²⁵ *Id.*

²²⁶ *Id.*

²²⁷ DAMASIO, *supra* note 154, at 36 ("The surgery was a success in every respect, and insofar as such tumors tend not to grow again, the outlook was excellent. What was to prove less felicitous was the turn in Elliot's personality. The changes, which began during his physical recovery, astonished family and friends. . . . In many ways . . . Elliot was no longer Elliot.").

²²⁸ *Id.* ("To be sure, Elliot's smarts and his ability to move about and use language were unscathed His knowledge base seemed to survive, and he could perform many separate actions as well as before.").

However, Elliot became disinhibited, undisciplined, and unemotive.²²⁹ In post-surgery tests, as Damasio wrote, “Elliot emerged as a man with a normal intellect who was unable to decide properly, especially when the decision involved personal and social matters.”²³⁰ Damasio gave Elliot a test that showed one additional post-operation change: Elliot’s VmPFC damage had compromised his ability to feel and process emotion, and thus to make personal and socially appropriate decisions.²³¹ “[T]he cold-bloodedness of Elliot’s reasoning,” writes Damasio, “prevented him from assigning different values to different options, and made his decision-making landscape hopelessly flat.”²³²

In a famous series of lesion studies using the Iowa gambling task (IGT),²³³ Bechara et al. examined the decision-making of patients with damage to the VmPFC.²³⁴ This set of studies suggested that VmPFC-damaged patients tend to be insensitive to future consequences of their choices, and are primarily guided by immediate gains.²³⁵ As Jeremy Gray observed, these findings suggest that patients with VmPFC damage “lack mechanisms of emotion-related feedback that healthy participants use to adaptively bias their choices in the IGT.”²³⁶ Therefore, VmPFC patients’ decision-making abilities are impaired likely due to “their lack of an appropriate affective basis on which to make adaptive choices.”²³⁷

²²⁹ *Id.* at 38 (“The tragedy of this otherwise healthy and intelligent man was that he was neither stupid nor ignorant, and yet he acted often as if he were. The machinery for his decision making was so flawed that he could no longer be an effective social being.”).

²³⁰ *Id.* at 43.

²³¹ *See id.* at 38–51.

²³² *Id.* at 51. *See also* Miller et al., *supra* note 224, at 102 (“Elliot’s problem lay in linking automatic valuation and regulatory control when necessary to make a decision.”).

²³³ Antoine Bechara et al., *Insensitivity to Future Consequences Following Damage to Human Prefrontal Cortex*, 50 *COGNITION* 7 (1994). The IGT is a psychological task designed to stimulate real-life decision making. Participants are required to choose between decks of cards that yield high immediate gain but larger future loss, i.e., a long-term loss, and decks that yield lower immediate gain but a smaller future loss, i.e., a long-term gain.

²³⁴ *Id.*; *see also* Antoine Bechara et al., *Characterization of the Decision-Making of Patients with Ventromedial Prefrontal Cortex Lesions*, 123 *BRAIN* 2189 (2000).

²³⁵ *See id.* at 2198 (“VM lesion patients preferred decks with high immediate reward to those with smaller reward, although the decks with small reward were more advantageous in the long term. VM lesion patients also preferred decks that had low immediate punishment to those with higher immediate punishment, although the decks with higher immediate punishment were more advantageous in the long run.”).

²³⁶ Jeremy Gray, *Affect and Action Control*, in *OXFORD HANDBOOK OF HUMAN ACTION* 277, 283 (Ezequiel Morsella, John Bargh & Peter Gollwitzer eds., 2009).

²³⁷ *Id.*

In another study, Bechara et al. compared the different roles of VmPFC and amygdala damages in volitional processes.²³⁸ They concluded that amygdala damages can hinder or strongly reduce one's ability to discern the emotional attributes of an emotionally charged stimulus.²³⁹ As a consequence, amygdala-damaged patients are unable to sufficiently "experience the emotional attributes of a stimulus that is charged with emotion."²⁴⁰ Similarly, a study by Hampton et al. indicated that patients with amygdala damage exhibited a profound change in PFC activity related to reward expectation and behavioral choice, indicating that information related to behavioral choice in PFC relies directly on input from the amygdala.²⁴¹ Altogether, these studies have indicated that abnormal functioning in the amygdala may well have a significant negative impact on emotional information processing at a cognitive level, resulting in poor control over behavioral responses.²⁴²

Additional support for this perspective emerges from studies on neuropsychiatric antisocial populations. These studies found links between deficits in socio-emotional brain circuits and inhibition.²⁴³ They suggested that those who show a lack of insight into their own behavior as well as a lack of moral emotions and emotional responses towards others (e.g., empathy)²⁴⁴ are less inhibited in violating the rights of others. For instance, injury to the VmPFC and the ACC—which, once again, are critical in emotional processing, as well as in behavioral motivation and regulation—have been linked to the onset of reckless and antisocial behavior without

²³⁸ Antoine Bechara et al., *Different Contributions of the Human Amygdala and Ventromedial Prefrontal Cortex to Decision-Making*, 19 J. NEUROSCI. 5473 (1999).

²³⁹ *Id.* at 5479 ("We see the impairment in decision-making after amygdala damage as an indirect consequence of the role of the amygdala in attaching affective attributes to stimuli.").

²⁴⁰ *Id.* at 5473.

²⁴¹ Alan N. Hampton et al., *Contributions of the Amygdala to Reward Expectancy and Choice Signals in Human Prefrontal Cortex*, 55 NEURON 545 (2007).

²⁴² See Steven Penney, *Impulse Control and Criminal Responsibility: Lessons from Neuroscience*, 35 INT'L J.L. PSYCHIATRY 99, 100 (2012).

²⁴³ See also Birgit Völlm, *Neurobiological Substrates of Antisocial and Borderline Personality Disorder: Preliminary Results of a Functional fMRI Study*, 14 CRIM. BEHAV. MENTAL HEALTH 39 (2004); Lau Siew Tee & Norshia Fauzan, *The Role of the Anterior Cingulate Cortex and Amygdala on Criminal Behavior*, 2 J. SCI. RES. & BEHAV. 203 (2015); Cole Korponay et al., *Impulsive-Antisocial Dimension of Psychopathy Linked to Enlargement and Abnormal Functional Connectivity of the Striatum*, 2 BIOL. PSYCHIATRY 149 (2017).

²⁴⁴ See also James Blair & Karina Blair, *Empathy, Morality, and Social Convention: Evidence from the Study of Psychopathy and Other Psychiatric Disorders*, in THE SOCIAL NEUROSCIENCE OF EMPATHY, Ch. 11 (2009); Richard Davidson et al., *Dysfunction in the Neural Circuit of Emotion Regulation*, 289 SCIENCE 591 (2000).

remorse.²⁴⁵ Importantly, as one study has highlighted, “[w]ithout the restraint of intuitive moral emotions and self-other conjoining . . . patients may not be able to deter an impulse to act in an unacceptable manner, even as they know right and wrong and understand the nature of their acts.”²⁴⁶

All things considered, neuroscience research impugns the legal understanding of self-control in two main respects. First, self-control is not a unitary capacity²⁴⁷ that either exists or does not exist. Rather, a multitude of complex and, sometimes, interrelated processes are involved self-control abilities. These processes involve brain mechanisms that are not purely or solely cognitive, but form complex circuits in which affective and motivational processes also play a prominent role in the evaluation, information processing, regulation, and appropriate reaction to salient stimuli. Thus, as has been observed, because abilities of self-control appear to depend on a multitude of dissociable processes, preserved functioning in one of these processes may still be accompanied by poor functioning in others.²⁴⁸ Although this does not imply a total lack of the capacity for self-control, this capacity may still be compromised.

Second, and consequently, people’s volitional power to choose which conduct to engage in within a given context, and thus to regulate their impulses, is not a logical consequence of their cognitive faculties of knowledge and understanding. The neuroscientific studies reported above suggest that the mechanisms involved in self-control support both cognitive and emotional processes, which serve distinct—yet related—functions in governing volitional faculties. A disruption in either cognitive or emotional processes (e.g., maladaptive emotional responses or deficient cognitive regulation) can *equally* endanger a given choice of appropriate behavior in response to a certain stimulus. Therefore, the factual or moral knowledge, or understanding, of the meaning of a certain action does not necessarily imply that antisocial impulses are controlled or that morally appropriate choices are made. As Jeffrey Rosen has asserted, “you can have a horrendously damaged brain where someone knows the difference between right and wrong but nonetheless can’t control their behavior. At that point,

²⁴⁵ See, e.g., Michael Koenigs, *The Role of Prefrontal Cortex in Psychopathy*, 23 REV. NEUROSCI. 253 (2012); Julian Motzkin et al., *Ventromedial Prefrontal Cortex is Critical for the Regulation of Amygdala Activity in Humans*, 77 BIOL. PSYCHIATRY 276 (2015).

²⁴⁶ Mendez, *supra* note 184, at 611.

²⁴⁷ See Buckholtz et al., *supra* note 212, at 13.

²⁴⁸ *Id.*

you're dealing with a broken machine, and concepts like punishment and evil and sin become utterly irrelevant."²⁴⁹

In sum, a rational moral agent appears to be one who retains sufficient capacities for choosing appropriate moral conduct, regardless of whether or not he or she is able to verbally determine whether that conduct is either right or wrong. Accordingly, not only should "self-control" and "knowledge" be evaluated and assessed separately, but the role that emotional faculties play in affecting self-control abilities should also be reconsidered.

IV. A TRIPARTITE TEST FOR LEGAL INSANITY

Neuroscientific insights into emotions and moral decision-making highlight that criminal law's ascriptions of the capacity for moral rationality to an alleged overriding cognitive sphere is limited and incomplete. As noted above, moral behavior is an integrated core of intertwined functions (i.e., cognitive and emotional) which only together may properly contribute to an individual's practical reasoning in accordance with the moral/normative requirements.²⁵⁰ It is flawed and unrealistic to maintain the view that higher cognition is the sole—or even the principal—controlling function of morally rational decision-making and behavior. Emotions and emotional processes significantly contribute to the mental equilibrium that allows individuals to behave in accordance with social, moral, and legal rules.

²⁴⁹ Jeffrey Rosen, *The Brain on the Stand*, N.Y. TIMES, (Mar. 11, 2007), <http://www.nytimes.com/2007/03/11/magazine/11Neurolaw.t.html?pagewanted=all> [<https://perma.cc/NQ8V-MNKG>].

²⁵⁰ See Patricia Greenspan, *Practical Reasoning and Emotion*, in THE OXFORD HANDBOOK OF RATIONALITY 206 (Alfred Mele & Piers Rawling eds., 2004) (arguing that the neuroscientific insights into emotions suggest that practical reasoning relies upon normal emotional development and functioning). Considering their guidance role, Greenspan holds that emotions are factors in our practical reasoning for two main reasons. First, emotions reinforce nonemotional reasons—e.g., desires, beliefs. By evaluating brute facts or stimuli through the attachment of positive or negative valences, emotions yield further non-emotional reasons "to sustain the conditions that make the evaluation appropriate." *Id.* This implies an understanding of emotions in normative terms, as providing or expressing reasons for actions. For instance, feeling moral guilt at the prospect of doing a given act can involve a negative moral judgment that the act is morally wrong along with an aversion to that act. This feeling of guilt might be an input to our desires and beliefs, and thus to our overall moral reasoning up to our behavioral outcome. Second, emotion provides *evaluative propositions*—e.g., that something might cause harm— and thus anticipate practical eventualities of actions or situations, thereby making an individual react accordingly. Therefore, when an actor chooses to do something for a reason, he does so on the basis of some sort of pro attitude toward actions of a certain kind, that is, some emotional commitment to it, whatever it is.

A reexamination of the role of emotions within the capacity for moral rationality entails an expansion of the relevant mental substance underlying the legal notion of insanity and produces two significant consequences for insanity tests. The first concerns the recognition of severe emotional deficits as eligible conditions for the insanity defense. The second concerns a different conception of volitional capacity as a multi-faceted construct involving a variety of functions (both cognitive and emotional): functions that are not necessarily related to the cognitive faculties of knowledge and understanding.

Altogether, such a reexamination would lead insanity tests to become tripartite and more dimensional, i.e., inclusive of cognitive, emotional, and volitional prongs. The cognitive prong would equate to an intellectual capacity test—that is, a test measuring the defendant’s knowledge or understanding of the factual and moral meaning of his or her conduct. The emotional prong would equate to an emotional capacity test—that is, a test measuring the defendant’s capacity to emotionally appreciate the moral significance of his or her conduct. The volitional prong would equate to a control test—that is, a test measuring the defendant’s capacity to control his or her impulses. Importantly, the volitional prong would be reconsidered to include emotion within its relevant application. Also, the volitional prong would be autonomous from the cognitive prong.

Expanding the substance of the volitional prong to incorporate emotional components also has consequences for the diminished capacity doctrine, as it is regulated by both the common law “heat of passion” and the MPC’s EED standards. In fact, rethinking the relationship between volition and emotion within insanity standards—that is, accepting that volitional impairments also meet insanity criteria when they are due to severe emotional dysfunctions—implies placing insanity and diminished capacity on a continuum, as the requirements to meet insanity and diminished-capacity standards would be the same. Therefore, the difference between insanity and diminished capacity proves to be purely quantitative, and diminished capacity is transformed into a “generic partial excuse” (or, perhaps more correctly, “generic partial insanity”), thereby integrating and supporting the validity of Professor Stephen Morse’s argument for introducing such a doctrine in Anglo-American criminal law.

A. LIMITING THE SCOPE OF THE COGNITIVE PRONG

As discussed, criminal law attributes the mental substance of insanity exclusively to cognition. Traditionally, the cognitive dimension is responsible both for agents’ cognitive faculties of understanding and knowing the factual and moral meaning of their conduct, and for their

volitional faculties of self-control as a consequence. Enriching the substance of the capacity for moral rationality with the contribution of emotion clearly implies the limitation of attributing such an absolute role to cognition within insanity tests.

Very simply, cognition remains the mental dimension that governs the evaluation in the cognitive prong of the insanity test: it remains responsible for the evaluation of one's capacity to propositionally understand the factual and moral meaning of his or her actions. The "factual knowledge" test remains one that measures whether an agent possessed a sufficient degree of cognitive or intellectual functioning to retain a factual understanding and knowledge of his or her action in a given context. For example, as some authors have suggested, the evaluation of this specific mental prong would encompass an agent's IQ level.²⁵¹ Regarding the "moral knowledge" test, because cognitive mechanisms are surely responsible for enabling agents' verbal understanding and knowledge of the moral significance of their actions (i.e., telling right from wrong), the new model of legal insanity retains the role of cognitive intelligence as mental source of agents' understanding and knowledge of the moral significance of their actions.

The difference with respect to the traditional model of legal insanity is that cognition is no longer treated as the sole mental faculty in charge of volitional faculties. This difference, of course, does not mean to deny that cognitive faculties contribute to one's capacity for self-control. As previously discussed, volitional processes involve their own mechanisms, both cognitive and emotional.²⁵² However, these mechanisms stand on their own; that is, they are not necessarily related to one's cognitive understanding and knowledge of the factual or moral significance of a given behavior, or its consequences. Thus, the "separation" of volition from cognition is simply intended to eliminate the "dependence" of the control test on the knowledge test, and thus to support a view of the volitional prong as an autonomous test.

B. INCLUDING AN EMOTIONAL PRONG

The empirical sciences reveal that people appreciate the social and moral significance of their acts, not only of their cognitive awareness that an act is wrong, but also of their experience of the usual moral feelings associated with that act.²⁵³ Critically, neuroscientific (and psychological)

²⁵¹ See *supra* note 215, at 31.

²⁵² See *supra* Part III.C.

²⁵³ See *supra* Part III.A.

studies indicate that possessing verbal or instrumental knowledge of something is not synonymous with being capable of moral decision-making and behavior if the relevant emotions are lacking.²⁵⁴ On the contrary, the existence of the former in the absence of the latter is a potential description of abnormal or even pathological conditions.²⁵⁵ Thus, emotional appreciation, in addition to intellectual awareness, should more correctly be viewed as an integral component of the capacities for moral rationality that form the mental preconditions of culpability and responsibility.²⁵⁶

Giving prominence to the emotional sphere within the capacity for moral rationality implies that an emotional prong would be explicitly included within an insanity ruling. The emotional prong of insanity tests would assess *agents' capacity to emotionally appreciate the moral significance of their actions*. Consequently, agents are potentially not culpable as long as they display such a disrupted emotional system that they are unable to perceive the rightness or wrongness of their actions, regardless of their verbal understanding or knowledge of the factual or moral facets of those actions. The emotional prong would thus measure whether the defendant in question possessed sufficient emotional capacities typically involved in moral judgments and behavior—for example, the capacity to feel empathic concern—in order to evaluate whether he or she was able to also perceive, or appreciate, the moral significance of the criminal act he or she committed.

To make these claims more solid, what follows examines which tangible effects the acceptance of emotional capacity as part of insanity tests could have on current insanity standards. Turning first to *M'Naghten*, the intellectual test it contains requires that a “defect of reason” arising from a “disease of [the] mind,” must impair the defendant’s ability to “know” the nature and quality of the act he or she was committing, as well as his or her capacity to know that the act was wrong.²⁵⁷ The inclusion of an emotional prong in the formulation and the substance of the *M'Naghten* rule would therefore have an impact on three different levels: first, the meaning of “reason”; second, the meaning of “disease of [the] mind”; and third, the meaning of “know.”

Let us begin with “reason.” According to traditional legal understandings, the concept of “reason”—as capacity for moral

²⁵⁴ *Id.*

²⁵⁵ *Id.*

²⁵⁶ See ST. JOHN'S L. REV., *supra* note 33, at 250 (“While an individual may understand both the nature of his act and its wrongfulness [cognition], he may nevertheless, due to mental illness . . . be so emotionally deranged [affection] as to be irresponsible.”).

²⁵⁷ *R. v. M'Naghten*, 8 Eng. Rep. 718 (1843).

rationality—encompasses only cognitive faculties. The discussion above, however, has suggested that the capacity for moral judgment and behavior also requires and depends on the normal functioning of emotional faculties, enabling an individual to also feel the moral significance of his or her actions. Thus, the defect-of-reason clause is no longer to be understood as an intellect-based defect of moral rationality but encompasses a broader notion of the capacity for moral rationality: one that also embraces emotional faculties.

If the scope of the word “reason” expands to also include emotional components, it follows that the concept of the sort of “disease of [the] mind” that can cause such a defect of reason must change and be expanded accordingly. That is, the disease-of-the-mind clause would also encompass severe emotional abnormalities as eligible conditions that may lead to a disruption of the capacity for moral rationality.

What is more, a broadening of both notions (“defect of reason” and “disease of [the] mind”) would inevitably lead to an expansion of the scope of the word “know” in a way that gives more weight to emotional components. However, to avoid conceptual and interpretive confusion, a more precise formulation of the test could maintain the knowledge requirement as an indicator in the cognitive test, but add an explicit provision for an emotional prong, which would require proof of the person’s lack of emotional appreciation of the moral significance of his or her criminal act.

Consequently, a potential reformulation of the *M’Naghten* test that considers the emotional prong reads as follows:

To establish a defense on the grounds of insanity, it must be clearly proven that, at the time of committing the act, the party accused was laboring under such a *defect of moral rationality*, from impairment of the party’s *cognitive or emotional capacities, or both*, of such extent as to lack the capacity to know the nature and quality of the act he or she was doing or *to appreciate emotionally* that the act was legally and morally wrong.

Regarding the ALI test, the current formulation seems to be more consistent with an emotionally informed understanding of the relevant capacities for moral rationality. In fact, the verb “to appreciate” seems to also contain an emotional test. As noted above, however, this term has caused much controversy because it is unclear how broadly it should be understood.²⁵⁸

Regardless of whether the verb “to appreciate” would be better interpreted more narrowly (as solely cognitive) or more broadly (as encompassing emotion), it can generate confusion. In fact, by attributing to

²⁵⁸ See discussion *supra* Part I.A.

the word “appreciate” both cognitive and emotional meaning, the line between these two forms of meaning could become excessively blurred. Ultimately, the risk would remain that relevance continues to be attributed solely to cognitive defects, without dedicating significant consideration to the emotional defects. Moreover, if one accepts that emotional (in)capacity plays a definite and autonomous role in moral judgments, it would appear to be more correct to split the word “appreciate” into two different requirements, such as *knowing and emotionally appreciating the criminality of one’s conduct*.

The provision of Section 4.01(1) would therefore read as follows:

A person is not responsible for criminal conduct if at the time of such conduct, as a result of a mental disease or disorder, *causing impairment of his or her cognitive or emotional capacities, or both*, he or she lacked substantial capacity either *to understand the factual and moral meaning of his or her conduct, or to appreciate emotionally the moral significance of his or her conduct*

Admittedly, introducing an explicit emotional prong in the first paragraph of the ALI test poses some challenges for the caveat paragraph contained in Section 4.01(2). Specifically, if emotional capacity becomes a prong of the insanity test, then the rationale of the caveat paragraph seemingly becomes meaningless. Admittedly, the categories of people to which the paragraph actually refers, namely psychopaths and more broadly people suffering from disorders that are characterized by severe socio-affective deficits, would become eligible for an insanity plea by virtue of the test’s newly introduced emotional prong.

The caveat paragraph, however, should not be eliminated and should continue serving its deterrent function, although to a different extent. Not all “repeated manifestations of criminal or otherwise antisocial conduct” are symptomatic of pathological emotional deficits,²⁵⁹ nor do emotional deficits necessarily lead to antisocial behavior. For emotional deficits to potentially exculpate defendants, it must be convincingly proven that they are sufficiently severely pathological to have a seriously compromised capacity for moral judgment at the time of the crime. Of course, it would be up to juries—with the help of expert witnesses—to evaluate and assess, case by case, the pathological seriousness of emotional derangements within the domain of insanity pleas. This sort of assessment is not new compared to what has been done to date with traditional tests.

²⁵⁹ *Wade v. United States*, 426 F.2d 64, 72–73 (9th Cir. 1970) (correctly pointing out that “[i]t is practically inconceivable that mental disease or defect would . . . be manifested only by repeated criminal or otherwise antisocial conduct”).

C. INTEGRATING EMOTIONS IN THE SUBSTANCE OF THE VOLITIONAL PRONG

The third feature of the new model of insanity concerns making the volitional prong (i.e., the capacity to choose which type of conduct to engage in and therefore to exert self-control) independent of the cognitive one. As observed above, the volitional prong in insanity tests plays a secondary role in comparison to the cognitive prong. Indeed, some insanity standards (e.g., the *M'Naghten* test) do not include an explicit control test at all, thereby placing the burden of insanity entirely on the cognitive knowledge test. Other tests, however, do require that agents are capable of controlling their impulses at the time of the crime and therefore also of conforming their behavior to what the law prescribes (according to the MPC model).²⁶⁰ However, even in this case, the autonomy of the volitional prong proves to be mostly theoretical and poorly applied.

As previously discussed, neither of these two conceptions of volitional incapacity explicitly consider the emotional dimension. The reason, as set out above, is that emotions are not considered to be part of what comprises the capacity for moral rationality. Therefore, even when emotions are impaired and provoke a lack of self-control, the law does not understand this condition to entail a lack of moral rationality that could serve as grounds for an excuse.²⁶¹ Ergo, when the lack of self-control depends on a cognitive defect—that is, when a person is not able to control his or her impulses and to conform his or her behavior due to an altered perception of reality—this condition constitutes reasonable grounds for an insanity defense.²⁶² However, when the lack of self-control is due to an emotional defect—i.e., when a person is not able to control his or her impulses because of an extreme (pathological) emotional disturbance—this condition is not considered to form acceptable grounds for an insanity defense.²⁶³ Rather, a lack of self-control due to emotional impairments alone at best constitutes a mitigating factor, one that could lead to the recognition of a lesser degree of culpability in terms of the degree of crime, as in the EED defense provided by the MPC.

As discussed above, self-control appears to rely upon a variety of distinct cognitive and socio-emotional processes, each of which contributes to individuals' control abilities.²⁶⁴ Importantly, preserved functioning in

²⁶⁰ MODEL PENAL CODE § 4.01 (AM. LAW INST. 1962).

²⁶¹ See *supra* Part II.B.

²⁶² See *supra* Part I.B.

²⁶³ See *supra* Part II.B.

²⁶⁴ See *supra* Part III.C.

one of these processes may still be accompanied by poor functioning in others.²⁶⁵ Furthermore, the processes involved in self-control may be irrespective of an individual's moral knowledge or understanding of his or her conduct. Thus, even if an individual is able to know right from wrong, he or she may still be unable—due to a disruption affecting self-control mechanisms (either cognitive or emotional)—to apply that knowledge to action.

If insanity standards were based on a definition that accepted this broader and multi-faceted vision of volitional capacity, then an incapacity for self-control would be allotted the same amount of relevance regardless of whether it is the result of severe emotional or cognitive defects. Thus, the substance of volitional incapacity could be reconceptualized as an incapacity to choose which conduct to engage in, resulting from deficits in either cognitive or emotional processes involved in self-control abilities. Importantly, this evaluation would be autonomous and irrespective of the cognitive faculties of knowledge and understanding.

The reconceptualization of the volitional prong as an autonomous prong would have diverse implications for different insanity standards. Let us begin with the *M'Naghten* rule. As previously discussed, the traditional formulation of this test does not encompass a volitional component. The reason, as elaborated above, lies in the fact that the test presumes cognition to be the only relevant mental dimension responsible for moral behavior, as it produces the knowledge of the nature and quality of an act, as well as of whether the act violates society's morals. If one is in possession of this knowledge, then he or she is presumed to be able to control any impulses.

Accepting that volitional processes have their own cognitive and emotional mechanisms, which are irrespective of one's cognitive faculties of knowledge and understanding, implies that the narrow cognitive test contained in the *M'Naghten* rule would be expanded to also include a volitional capacity test. A potential re-formulation of the test with the introduction of a volitional prong, in addition to a cognitive and an emotional one, reads as follows:

To establish a defense on the grounds of insanity, it must be clearly proved that, at the time of committing the act, the party accused was laboring under a *defect of moral rationality, from impairment of the party's cognitive and/or emotional capacities, or both*, of such extent as to lack the capacity to know the nature and quality of his or her act, to appreciate emotionally that the act was wrong, or *to conform his or her behavior to his or her knowledge or appreciation*.

²⁶⁵ *Id.*

The same mechanism can be applied to the ALI test. Reconsidering the control test contained in Section 4.01(1) of the MPC from the perspective of an understanding of volitional capacity that also depends on emotional factors, this provision can be reformulated as follows:

A person is not responsible for criminal conduct if at the time of such conduct as a result of a mental disease or disorder, *causing impairment of his or her cognitive or emotional capacities, or both*, he or she lacked substantial capacity . . . to control his or her conduct in the circumstances and thus to conform his or her conduct to the requirements of the law.

As this reformulation makes clear, the introduction of an emotion-oriented notion of volitional (in)capacity would have no substantial impact on the ALI test, in the sense that it would simply specify the substance of the control test without significantly altering the test's original formulation. The new formulation proposed here would simply make the test more specific, so as to overcome the conceptual imprecisions that the current text contains and to ensure there less space for interpretive uncertainty.

D. RETHINKING DIMINISHED CAPACITY AS GENERIC PARTIAL INSANITY

In addition to affecting the insanity doctrine, an emotion-oriented paradigm of the capacity for moral rationality may have significant repercussions for the diminished-capacity doctrine—such as the common law's "heat of passion," the provocation or passion doctrine, or EED doctrine adopted by the MPC and several single statutes in the US.

Let us briefly recapitulate on both the provocation or passion and the EED doctrines.²⁶⁶ Both doctrines are substantially characterized by a lack of control due to an emotional breakdown at the time of the crime, whereby the former is due to a provocation, and the latter is due to an EED for which there is a reasonable explanation or excuse, regardless of whether there was any previous provocation. The rationale underlying both doctrines is that a temporary lapse of control due to an intense emotional breakdown reduces, albeit temporarily, an individual's rationality at the time of the crime.²⁶⁷ Even so, the law does not treat this type of lack of self-control as an actual moral rationality defect. In fact, the scope of such mitigations is to lessen

²⁶⁶ See *supra* Part II.B.

²⁶⁷ *State v. Gounagias*, 153 P. 9, 11–12 (Wash. 1915) ("The doctrine of mitigation is briefly this: That if the act of killing, though intentional, be committed under the influence of sudden, intense anger, or heat of blood, obscuring the reason, produced by an adequate or reasonable provocation, and before sufficient time has elapsed for the blood to cool and reason to reassert itself.").

only the degree of the crime, not the perpetrator's overall culpability.²⁶⁸ Also (and herein lies another paradox), both doctrines are limited to reducing murder to manslaughter, and neither can be applied to crimes in general.

The contradictory nature of the diminished-capacity doctrine has given rise to several critiques. One compelling argument against the current state of the diminished-capacity doctrine has been offered by Stephen Morse,²⁶⁹ who vigorously affirms that diminished capacity should take on broader connotations. Morse acknowledges that, contrary to the "all-or-nothing doctrines" adopted by current criminal law, capacities for moral rationality, including self-control, are continuum concepts.²⁷⁰ In view of the fact that people "display an enormously wide range of rational and control capacities,"²⁷¹ a truly fair judgment of culpability and responsibility, as well as a fair determination of punishment, must consider the kinds of impairments that affect one's rationality to some significant degree, even if they do not entirely compromise it.²⁷² Furthermore, Morse rightly indicates that limiting the diminished-capacity doctrine to homicide alone is pointless, considering that any crime can be committed by a defendant whose rational capacities are to some degree impaired.²⁷³ The applicability of the mitigating factors in the present EED and heat of passion defenses should be extended to cover all crimes. With this in mind, Morse proposes the adoption of an additional verdict, namely the "Guilty but Partially Responsible" verdict, as a new general affirmative defense that "requires a substantial diminution in rationality because less serious impairments are sufficient to warrant lesser blame and punishment."²⁷⁴

Morse's argument remains intentionally vague about the kinds of mental impairments that may meet the requirements of the generic partial excuse he proposes.²⁷⁵ He argues that the generic partial responsibility excuse should apply in cases of less severe rationality defects, which implies cognitive and—irrespective of whether they are clearly provided for

²⁶⁸ See *supra* Part II.B.

²⁶⁹ Morse, *supra* note 129.

²⁷⁰ *Id.* at 296.

²⁷¹ *Id.*

²⁷² *Id.* at 302 ("Perhaps the law should adopt a generic mitigation (that would consider degrees of rationality and responsibility diminution).").

²⁷³ *Id.* at 296 ("Compromised rationality and its effect on culpability are not limited to homicide. Fairness and proportionality require that doctrinal mitigation should be available in all cases in which culpability is substantially reduced.").

²⁷⁴ *Id.* at 299–304.

²⁷⁵ *Id.* at 301 ("I would trust legislative judgment or the common law process to identify which rationality-diminishing factors would be justified.").

in the text—self-control defects.²⁷⁶ These are basically the same kinds of defects (though in a lesser degree) that warrant a designation of insanity.

Morse also argues that the partial generic excuse he proposes would have the effect of extending the applicability of the heat of passion and EED doctrines to all crimes, thereby implying that loss of self-control arising from emotional disturbance should also be considered a rationality defect.²⁷⁷ In so arguing, Morse seems to implicitly attribute relevance to the emotional factor in matters of self-control and, more broadly, rationality.

If this interpretation is correct, then Morse's theory of diminished capacity lends support to the emotion-oriented descriptions of moral rationality and insanity proposed here. In fact, building on an emotion-enriched view of the capacity for moral rationality, and consequently of self-control, the net distinction between the kinds of mental impairments necessary to warrant insanity and those needed to warrant diminished capacity prove meaningless. Once one recognizes that emotions also play a role in the capacity for moral rationality, including self-control, there is no reason to maintain a doctrinal distinction between volitional impairments due to cognitive defects and volitional impairments due to emotional defects. Rather, by making emotions an integral part of the capacity for moral rationality that grounds the insanity defense, it follows that the kinds of requirements that ground both doctrines are exactly the same, though to different degrees. That is, diminished capacity would become a concept *in continuum* with insanity.

Furthermore, as Morse correctly points out, a generic partial excuse should apply in all cases of partial rationality defects. In combining this application of the generic partial excuse with the updated notion of the capacity for moral rationality, also including an emotional prong, it follows that a generic partial excuse would not apply only in cases in which a defendant's mental conditions impair to some degree his or her cognitive

²⁷⁶ *Id.* at 295–98 (“I claim that the best interpretation of our moral and criminal law excusing practices is that there are only two basic excusing conditions: diminished rationality and ‘hard choice.’ . . . [T]here is a limited need for an excuse based on an impaired capacity for self-control. The capacity for rationality, the ‘hardness’ of choice, and the capacity for control are all continuum concepts. Nonetheless, with precious few exceptions, present criminal law contains doctrinal all-or-nothing, bright line tests Present law is unfair because it does not sufficiently permit mitigating claims The solution to all these problems of potential doctrinal deformation is a generic mitigating excuse.”).

²⁷⁷ *Id.* at 295 (“Provocation/passion and extreme mental or emotional disturbance as partially excusing mitigating doctrines are best explained by the theory that these conditions non-culpably reduce the capacity for rationality.”).

and volitional faculties, but also when these conditions compromise his or her emotional capacities. This means that the mental diseases or disorders that could lead to a total insanity verdict could also lead to a generic partial excuse or—perhaps more correctly—a partial insanity verdict, depending on their intensity and the degree to which the defendant’s cognitive, emotional, and volitional capacities are compromised. Thus, the substance of the generic partial insanity tests would be tailored to that of “total” insanity tests, with any difference between these two doctrines being merely quantitative, not qualitative. Importantly, turning the diminished-capacity doctrine (as it currently stands) into a partial-insanity doctrine further confirms the pointlessness of limiting the application of the diminished-capacity doctrine to homicide.

On a final note, the provision for a partial insanity doctrine raises two sets of normative issues. First, the law should determine what degree of impairment is needed to fall within either of the two forms of insanity. Second, the law should determine what kind of sentence could follow a partial insanity verdict. These issues, though critical, go beyond the scope of this article. Further research is needed to explore them in more detail and to set a hypothetical normative framework that regulates the implementation of this newly introduced partial excuse.

V. NORMATIVE ARGUMENTS SUPPORTING AN EMOTION-ORIENTED MODEL OF LEGAL INSANITY

The model of legal insanity proposed here is vulnerable to criticism. Admittedly, “[u]neasiness about science’s interference with legal understandings of cognition and responsibility, along with law’s outmoded view of mental illness, work concurrently to frustrate the progress of [the insanity] defense based on neurobiological evidence,” as one author wrote.²⁷⁸ Also, attempts to root the substance of the insanity defense in more scientific grounds are not new to psychiatry or legal scholarship. (Neuro)psychiatry and the (neuro)behavioral sciences in general have adopted and suggested a far-reaching understanding of mental disease, one that also encompasses affective and volitional disorders.²⁷⁹

²⁷⁸ Jozsef Meszaos, *Achieving the Peace of Mind: The Benefits of Neurobiological Evidence for Battered Women Defendants*, 23 *YALE J.L. & FEMIN.* 117, 172 (2011).

²⁷⁹ See, e.g., AMERICAN PSYCHIATRIC ASSOCIATION, *DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDERS 20* (5th ed. 2013) (“A mental disorder is a syndrome characterized by clinically significant disturbance in an individual’s cognition, emotion regulation, or behavior that reflects a dysfunction in the psychological, biological, or developmental processes underlying mental functioning. Mental disorders are usually associated with significant distress or disability in social, occupational, or other important

Notwithstanding, criminal law has intentionally remained mired in an intellectualist model of mental normalcy and therefore also of mental abnormality.²⁸⁰

In addition to the historical reasons illustrated above, the rationale underlying this position is especially one of criminal justice policy.²⁸¹ Indeed, the currently dominant intellectualistic conception of insanity is intended to meet the retributive,²⁸² deterrent,²⁸³ social-control, and social-security²⁸⁴ needs of criminal justice. While I comprehend the legal and crimino-political arguments for keeping a narrow model of legal insanity, these reasons are neither absolute nor insurmountable. There are at least three normative counterarguments that can be offered in support of an expanded model of insanity that takes into account emotional components: 1) an emotion-oriented model of legal insanity is in greater compliance with the principle of personal guilt; 2) an emotion-oriented model of legal insanity is more able to grasp the essence of blameworthiness and just

activities. An expectable or culturally approved response to a common stressor or loss, such as the death of a loved one, is not a mental disorder.”).

²⁸⁰ See Dillard S. Gardner, *Insanity as a Defense in the North Carolina Criminal Law*, 30 N.C. L. REV. 4, 7 (1951) (“For example, the notion that intelligence may be separated from the volitional and emotional life of an individual appears absurd and fantastic to most scientists, but appears to be taken for granted in legal theory.”).

²⁸¹ Julie E. Grachek, *The Insanity Defense in the Twenty-First Century: How Recent United States Supreme Court Case Law Can Improve the System*, 81 IND. L.J. 1479, 1481–82 (2006) (“The insanity defense addresses the policy issues inherent in the question of criminal culpability Society’s recognition of a moral difference between the acts of a sane person and a mentally ill person results in the insanity defense serving dual roles in the criminal justice system: 1) as a way to distinguish between offenders who are able to conform their conduct to the law as a result of punishment from those offenders who are not able to conform their conduct to the law despite punishment, and 2) as a method of ensuring offenders posing a threat to society are restrained.”).

²⁸² Stephen J. Morse & Richard J. Bonnie, *Abolishing the Insanity Defense Violates Due Process*, 41 J. AM. ACAD. PSYCHIATRY L. 488, 489 (2013) (“In the criminal justice system, an offender who lacks the capacity to understand the wrongfulness of his actions as the result of severe mental disorder does not deserve full blame and punishment.”).

²⁸³ *Id.* (“[O]ffenders cannot be appropriately deterred because the rules of law and morality cannot adequately guide them.”).

²⁸⁴ Wallace A. MacBain, *Insanity Defense: Conceptual Confusion and the Erosion of Fairness*, 67 MARQ. L. REV. 1, 16–17 (1983) (“[M]an possesses a free will, a capacity to make rational choices for which he should be held accountable . . . mentally abnormal offenders, determined to be nonresponsible and thus beyond the reach of criminal law, will be subjected to alternative rehabilitation and preventative mechanisms to assure adequate social control . . . one appropriate function of the insanity defense is the identification and exculpation of one whose capacity for free choice has been so diminished as to warrant a characterization of nonresponsibility with the consequent invocation of alternative social control mechanisms.”).

deserts that lie at the core of retributive punishment; and 3) an emotion-oriented model of legal insanity strengthens the aims of rehabilitation.

A. PERSONAL GUILT

In “Law and Psychiatry,” Michael Moore claims that criminal law does and should settle for its folk psychological descriptions of human thought—even though they may be limited—because that is all the law needs to hold someone criminally responsible.²⁸⁵ As he writes:

The very abstract view of persons in terms of autonomy and rationality is of course radically incomplete as a picture of any person we know. In particular, left out is the life of the emotions where, if anywhere, the ‘affection of other men’ is gained. Yet the radical incompleteness of the law’s view of a person is no argument that it is *wrong*. As far as it goes, the law’s view of persons could be quite correct even if radically incomplete.²⁸⁶

Moore’s claim that the legal approach may remain true although it is radically incomplete exposes itself to criticism. This is because the incompleteness of the practical reasoning and decision-making conditions underlying the traditional conception of culpability and culpability doctrines is precisely a means of evading the truth regarding how individuals reason and make decisions in moral contexts, such as decisions concerned with offending.

Moving from this claim, the first argument for rethinking insanity with the aid of neuroscience draws on the normative consideration that criminal law’s reduction of blameworthiness to the dimension of cognition contrasts with the universal *principle of personal guilt*. Here, the word “personal” does not denote only that culpability can be attributed only to the actual perpetrator of a given criminal wrongdoing. Rather, it also indicates that culpability needs to be based on an individual’s actual mental and moral participation in the commission of an offence, that is, he or she internally approves of the act and decides to act in breach of legally protected values. Slightly differently formulated, a criminal act must belong to the mental and moral domain of the perpetrator, and must express his or her disregard, or lack of concern, for the interests of other individuals protected by the law. Ultimately, this is what makes a given unlawful behavior blameworthy.

To comply with the principle of personal guilt, the notion of culpability must embrace *all* relevant mental factors that contribute to an individual’s moral decision to act unlawfully. In this respect, I agree fully with Kimberly Ferzan’s proposal for a “holistic” understanding of

²⁸⁵ MOORE, *supra* note 29, at 112.

²⁸⁶ *Id.*

culpability, according to which culpability is understood as the *sum* of practical reasoning and decision-making conditions underlying any blameworthy state of mind.²⁸⁷ As Ferzan asserts, “[it is precisely] the sum of these parts that gives rise to our normative judgment about whether the actor’s reasoning gave due regard to the interests of others [i.e., whether it is blameworthy or not].”²⁸⁸ While Ferzan’s holistic account of culpability is very accurate in grasping the essence of the notion of personal guilt, it should be noted that the descriptive, cognition-based model of culpability (and of culpability-related doctrines, such as insanity), is not sufficient to effectively depict an individual’s attitude of disregard for legally protected values—that is, his or her blameworthiness.

In view of the illustrated (neuro)scientific teachings about moral judgments and antisocial behavior, it is clear that the decision to engage in moral types of behaviors—such as criminal behavior²⁸⁹—involves a far more complex mechanism in which emotional factors, and not only the cognitive ones, play a critical role. Therefore, the blameworthy essence of one’s decision to act against legally protected values cannot only be found in neutral cognitive states, but also and more extensively in the affective mechanisms that underlie and drive an individual’s judgment to opt for immoral, rather than moral conduct.

If culpability is viewed as the sum of the agent’s practical reasoning and decision-making conditions leading up to criminal actions, and since emotions *are* factors in the practical reasoning and decision-making conditions leading up to such actions, it follows that a notion of culpability that truly complies with the principle of personal guilt should also require that individuals possess sufficient emotional soundness at the time of the offence.

B. CULPABILITY-BASED RETRIBUTION

Retribution essentially punishes perpetrators for two reasons: the first, which is objective, is that the perpetrators do wrong acts (harm-based retribution); the second, which is subjective, is that the perpetrators know that the acts they do are wrong, and yet choose to act upon their antisocial

²⁸⁷ Kimberly Kessler Ferzan, *Holistic Culpability*, 28 CARDOZO L. REV. 2523, 2523 (2007).

²⁸⁸ *Id.* at 2544.

²⁸⁹ See Per-Olof H. Wilkström & Kyle Treiber, *The Role of Self-Control in Crime Causation*, 4 EUR. J. CRIMINOLOGY 237, 244–45 (2007) (describing crimes as “moral actions,” i.e. actions that follow or break moral rules, and hence are “guided by what is right and what is wrong to do”).

impulses (culpability-based retribution).²⁹⁰ Central to any retributive punishment is the premise of moral blameworthiness, which presupposes moral rationality.²⁹¹ Thus, the perpetrators ultimately deserve punishment precisely because they are rational individuals who consciously choose to act immorally and do harm.

The understanding of moral rationality that underlies and justifies retributive punishment is grounded in the same cognition-based psychological assumptions about the mental ingredients of moral rationality as those that ground culpability.²⁹² Only an individual with a solid intellectual capacity is assumed to be able to make a conscious decision to engage in conduct that rejects his or her community's moral norms and, consequently, to appreciate the retributive force of the punishment for his or her misconduct.²⁹³ On the contrary, it makes no sense to punish someone who is not able to comprehend the retributive path of punishment. From this perspective, retribution is the logical conclusion of a normative syllogism, the premise of which is precisely a cognition-based view of moral rationality and hence of culpability. If culpability presupposes intellect-based moral rationality, and retribution presupposes culpability, then punishment is justified as long as it is directed to individuals who are provided with the intellectual faculties that make them morally rational.

If the premises of this syllogism are changed, the conclusions inevitably also change. In fact, if the substance of the capacity for moral rationality, and hence of culpability, is enriched with certain emotional faculties, it follows that punishment is justified only when these conditions are also met. Differently said, if being an individual capable of moral rationality, and consequently of moral blameworthiness, also requires possessing specific emotional capacities, then retributive punishment is justified as long as it is addressed to people who do possess these capacities. Because mentally diseased or disordered individuals cannot justifiably be punished for their acts on the grounds that they do not possess

²⁹⁰ See Beatrice R. Maidman, *The Legal Insanity Defense: Transforming the Legal Theory into a Medical Standard*, 96 B.U. L. REV. 1831, 1843–44 (2016).

²⁹¹ MOORE, *supra* note 29, at 244.

²⁹² Richard L. Lippke, *Retribution and Incarceration*, 17 PUB. AFF. QUART. 29, 43 (“Retributivism insists that legal punishment be structured so that, at a minimum, it is consistent with treating offenders as moral beings capable of understanding the wrongs they have committed and the fairness of the penal sanctions imposed on them by the state in response to those wrongs.”).

²⁹³ See Donald A. Dripps, *Fundamental Retribution Error: Criminal Justice and the Social Psychology of Blame*, 56 VAND. L. REV. 1383, 1424 (2003) (observing that “culpability-based retributivists believe that the actor’s subjective awareness of wrongdoing triggers blameworthiness and makes the actor eligible for punishment”).

the relevant capacities for moral rationality—including the emotional capacities—then a legal insanity defense has to be sufficiently broad to encompass people whose mental illnesses rendered them unable to be morally rational at the time of the crime.

In addition to these theoretical considerations, an expanded notion of legal insanity would also solve one of the main practical paradoxes of the criminal justice system, namely punishing mentally ill individuals with (often severe) socio-emotional impairments on the sole grounds that they do retain a cognitive capacity to know, or to understand, what they did at the time of the crime and that what they did was wrong. The practical and paradoxical result is that these people get sentenced to a given penalty (let us assume medium- to long-term imprisonment) without actually being able, due to their mental illness, to perceive, appreciate, or follow the “re-educative” path that retributive punishment is intended to incite. In addition, conventional punishment significantly undermines their successful reentry into the community and social reintegration subsequent to serving a (prison) sentence.²⁹⁴

C. REHABILITATION

The model of insanity proposed here would also better suit the purposes of rehabilitation. Rehabilitation presupposes that punishment is justified as long as it can provide individuals with the means to be socially functional and, thus, to return to society as law-abiding citizens.²⁹⁵ To accomplish the goals of rehabilitation, adjudication and sentencing must be as careful and individualized as possible; that is, they must be tailored to the actual needs of individual perpetrators.²⁹⁶ The inclusion of emotional capacities in the legal notion of insanity would definitely allow for a more accurate evaluation of an individual’s personality in relation to the crime committed. It may also offer a defendant who suffers from a severe mental illness (aside from cognitive diseases)²⁹⁷ that appears to have affected his or

²⁹⁴ See Arthur J. Lurigio et al., *The Effects of Serious Mental Illness on Offender Reentry*, 68 FED. PROB. 45 (2004), available at http://www.uscourts.gov/sites/default/files/68_2_9_0.pdf [https://perma.cc/XUU2-BY6Q].

²⁹⁵ See *Rehabilitation*, in PRINCIPLED SENTENCING: READINGS ON THEORY AND POLICY 1 (Andreas von Hirsch et al. eds., 3d ed. 2009).

²⁹⁶ *Id.*

²⁹⁷ The adoption of a tripartite insanity test implies a broadening of the category of individuals that can potentially fall within the spectrum of this defense. The updated legal notion of mental illness would also include mental diseases or disorders characterized by a pathological lack of emotions and moral feelings towards other subjects, as well as those characterized by severe impairments in behavioral control—as long as it is convincingly proved that these conditions significantly compromised the person’s ability to refrain from

her capacity for prosocial behavior to attain access to rehabilitation programs that ordinary incarceration does not provide.

As previously noted, even in those cases in which a defendant's psychiatric diagnosis is one of a mental disease or disorder that compromises his or her socio-affective faculties, this condition may nonetheless prove insufficient to meet the legal criteria of insanity, and the defendant would therefore still be considered entirely punishable. From a rehabilitative standpoint, as an abundance of literature shows, the result of punishing these people is the worsening of their socially vulnerable personalities.²⁹⁸

At this point, critics could raise two main objections. Foremost, if current approaches to indefinite commitment of people found "not guilty by reason of insanity" (NGRI) are continued, prisoners who once would have been eligible for release after a determinate period of time could now be held for the rest of their lives—not necessarily something they would consider a positive development. The second objection is that people with specific socio-affective disorders, especially psychopathy, are generally resistant to clinical treatment and have a proclivity to engage in stable criminal behavior, that is, they are socially dangerous.²⁹⁹ Admittedly, there is still little evidence regarding how this class of individuals can be successfully and appropriately treated.³⁰⁰

However reasonable in a first gloss, these practical objections are not insurmountable. Civil libertarian concerns about the risks of rerouting perpetrators with severely pathological socio-affective problems from the prison system to the civil commitment system fail to consider the point made above; that is, conventional incarceration has detrimental effects

antisocial behavior. See George B. Palermo, SEVERE PERSONALITY-DISORDERED DEFENDANTS AND THE INSANITY PLEA IN THE UNITED STATES: A PROPOSAL FOR CHANGE (2010), quoted by Ralph Slovenko, *Commentary: Personality Disorders and Criminal Law*, 37 J. AM. ACAD. PSYCHIATRY L. 182, 183 (2009) ("Individuals who have a severe personality disorder should be allowed to enter a plea of total or partial insanity based on evidence of a decompensation into irrational behavior at the time of the alleged crime, and should be allowed to present all exculpatory evidence available to them to prove their claim.").

²⁹⁸ For an excellent empirical and normative analysis of the actual negative impact and the concrete risks of imprisonment when dealing with prisoners with mental disorders, see E. Lea Johnston, *Vulnerability and Just Desert: A Theory of Sentencing and Mental Illness*, 103(1) J. CRIM. L. & CRIMINOLOGY 147 (2013).

²⁹⁹ See, e.g., James R. P. Ogloff & Melisa Wood, *The Treatment of Psychopathy: Clinical Nihilism or Steps in the Right Direction?*, in RESPONSIBILITY AND PSYCHOPATHY: INTERFACING LAW, PSYCHIATRY, AND PHILOSOPHY 155 (Luca Malatesti & John McMillan eds., 2010).

³⁰⁰ *Id.*

especially on this class of individuals.³⁰¹ Although prison sentences last for determinate periods of time, when these individuals are confined to a prison for their sentence, they are released with the same (or even worsened) mental conditions as they had when they began their sentence.³⁰² In contrast, if these individuals were rerouted to (an ideally reformed)³⁰³ civil commitment system, there would be the possibility of dealing with them more effectively than with conventional incarceration. Ideally, as soon as they are assessed to no longer pose a threat to themselves and to others, they would be released with less propensity to engage in socially dysfunctional behavior as when they began their rehabilitation program. One option may be the construction or the staffing of specialized rehabilitation centers—using James Gilligan’s definition, “residential communities” or “centers for human development”³⁰⁴—as well as a greater

³⁰¹ See, e.g., Arielle R. Baskin-Sommers & Karelle Fonteneau, *Correctional Change Through Neuroscience*, 85 *FORDHAM L. REV.* 423, 428 (2016) (arguing that the socially scarce conditions of a prison setting will most likely generate or exacerbate “neurobiological deficits and maladaptive behaviors This becomes a significant issue, especially for individuals who are chronic offenders, where existing neurobiological vulnerabilities are intensified in settings of confinement and segregation, thereby reinforcing maladaptive patterns of behavior”); see also James Bonta & Paul Gendreau, *Reexamining the Cruel and Unusual Punishment of Prison Life*, 14 *LAW & HUM. BEHAV.* 347 (1990).

³⁰² See *Wade v. United States*, 426 F.2d 64, 72 (moving a criticism to the caveat paragraph of the ALI test, the court points out that the exclusion of persons that are diagnosed as psychopathic but are nonetheless “seriously ill and . . . incapable of persistent, ordered living of any kind” leaves society unprotected, since after serving their prison sentence they would be released with the same issues they had at the beginning of their sentence).

³⁰³ As many authors have correctly urged, the civil commitment system should be profoundly reformed. Current involuntary hospitalization should be replaced with alternative, less stigmatizing, and less harsh therapeutic approaches, in the most humane and community-like settings as possible. See, e.g., Richard Bonnie, *Reforming Civil Commitment: Serving Consumers’ Needs While Protecting Their Rights*, *MENTAL HEALTH L.* 3 (2006); Stephen J. Morse, *Mental Disorder and Criminal Justice*, in *REFORMING CRIMINAL JUSTICE: A REPORT OF THE ACADEMY FOR JUSTICE BRIDGING THE GAP BETWEEN SCHOLARSHIP AND REFORM 251*, 320–23, (Erik Luna ed., 2018)

³⁰⁴ See James Gilligan, *A Modest Proposal to Universalize the Insanity Defense and Replace Prisons and Punishment with Treatment and Education*, 12 *INT’L J. APPLIED PSYCHOANAL. STUD.* 134, 147 (2015) (“What I am recommending in this article is . . . that we create residential communities that will be more humane and more therapeutic than our old mental hospitals were, and more than our prisons and jails are today, and adapt them to the treatment and education not only of those with the major ‘Axis I’ disorders, but also of those with the major personality disorders that are the main causes of serious or life-threatening interpersonal violence. . . I would propose that we create an entirely new type of secure residential college and therapeutic community that could function as a human development center, in which to enable those . . . who have suffered [and caused others to suffer] from the physical, emotional and cognitive abuse and deprivation that had prevented

implementation of outpatient rehabilitation programs when conditions allow.

Regarding the second objection, the lack of unanimous consensus about effective clinical treatments for offenders with certain types of mental disorders—like psychopathy—does not mean that the conventional prison system is the only or the right response to this particularly thorny class of individuals, nor that research should lose sight of identifying alternative ways to deal with them. For instance, the British neuroscientist Daniel Reisel hypothesizes that empathy training programs in secure, socially-stimulating environments other than prison facilities could be an asset to treat offenders with socio-affective deficits, such as psychopaths have.³⁰⁵ He posits that relationally-based situations could prove the most suitable way to create new opportunities for neural growth in the emotional circuits of the brain, increasing the likelihood of fostering empathy and sociable tendencies and enabling these offenders' positive emotional transformations towards pro-social attitudes.³⁰⁶ These insights could lead to an increased adoption of individualized, inclusionary socio-rehabilitative measures, which act as positive incentives for high-risk offenders' emotional healing and social functioning, thereby reducing risks of recidivism. Admittedly, very little is known about this avenue of treatment. Ensuring that research on these alternative treatment options remains a central tenet of future initiatives is thus crucial.³⁰⁷

them from achieving the degree of healthy, life-sustaining development and maturation—for example, development of the emotional capacities for love, care, concern, empathy, and a sense of responsibility for the welfare of others . . .”). See also Barbara Dickey et al., *Therapeutic Communities and Mental Health System Reform*, 32 PSYCHIATRIC REHABILITATION. J. 105 (2008) (describing the value of therapeutic communities to deal with people with serious mental illness).

³⁰⁵ Daniel Reisel, *The Neuroscience of Restorative Justice*, TED TALK (Mar. 18, 2014), <https://www.youtube.com/watch?v=tzJYY2p0QIc> [<https://perma.cc/K9C4-XUB8>].

³⁰⁶ *Id.*; see also ANDREA L. GLENN & ADRIAN RAINE, PSYCHOPATHY: AN INTRODUCTION TO BIOLOGICAL FINDINGS AND THEIR IMPLICATIONS 170 (2014) (suggesting “[h]ousing [offenders who are at continuous high risk of committing serious offenses] in a location that is secure but that practices humane treatment, minimizing aspects of punishment and allowing the individuals to have as much freedom as possible given the constraints of keeping such offenders away from society” and proposing “an increased focus on developing new treatment programs”).

³⁰⁷ See, e.g., the ‘Dangerous and Severe Personality Disorder Program (DSPD),’ which was launched by the UK Government and aimed at dealing with offenders suffering from personality disorders who posed a significant risk of harm to others and themselves. Although this arguable initiative was harshly criticized on several grounds—for instance, it lacked clinical evidence to treat personality disorders, and it allowed preventive detention for people who had not been convicted of an offence—many authors acknowledge that it changed political attitudes toward offenders with personality disorders. As Ruth Scally

CONCLUSION

This article offers a potential approach to how neuroscientific findings may be used to reshape the doctrine of legal insanity. Its ultimate goal is not to provide definitive answers to the issues surrounding the insanity defense, but simply to lay some grounds on which to initiate further debates in criminal law. Further research is needed to explore the practical corollaries of this new model of legal insanity. For instance, it would be necessary to investigate how the newly introduced emotional prong of insanity standards would be assessed through reliable judgments (both empirical and normative). Notably, it would be necessary to set parameters to determine when socio-affective deficits or abnormalities (assessed with behavioral and, perhaps, neuroscientific measures) could actually support a judgment of a defendant's lack of, or severe impairment of, his or her emotional capacity in relation to the crime committed. Another line of research could investigate in more detail the dividing line between total and partial insanity, as well as the sentencing regime following a hypothetical generic partial insanity verdict. Last, but not least, future research could enquire after the possible implications of this model of insanity (both total and partial) for contentious classes of offenders. The most notable example would be offenders with severe psychopathy. These tasks are by no means easy and require careful and detailed elaboration at an interdisciplinary level. Yet practical difficulties should not be used as a reason to not imagine circumstances in a different light.

observes, although the program no longer exists, "the needs of personality disordered offenders have remained on political agenda, as demonstrated by the continued investment in the offender pathway currently being implemented." Ruth Scally, *The DSPD Programme: What Did It Tell Us About the Future for Managing Dangerous Prisoners with Severe Personality Disorders?*, in *MENTAL HEALTH, CRIME AND CRIMINAL JUSTICE: RESPONSES AND REFORMS* 184, 194 (Jane Winstone ed., 2016).

