LINCOLN MEMORIAL UNIVERSITY
LAW REVIEW

VOLUME 10 FALL 2022 ISSUE 1

FORMING A "BRAIN PRINT:" USING COGNITIVE NEUROSCIENCE AND BRAIN IMAGING AS AN OBJECTIVE MEASURE OF CRIMINAL INSANITY

By Christos D. Strubakos Ph.D., J.D. (Cand. 2023)

"Yes, that's what it was! I wanted to become Napoleon, that is why I killed her…Do you understand now?" – Fyodor Dostoyevsky, *Crime and Punishment.*

I. INTRODUCTION[1]

Imagine a scenario where police arrest a defendant for domestic violence. When the arresting officers arrive at the scene and detain him, he utters a threat in perfect English, stating he will return home and kill his wife upon his release from jail. He is swiftly taken off the premises and booked at the jail. While waiting for his bond hearing, the accused suffers what later becomes characterized as a psychotic break. He urinates on the floor and smears his feces throughout his cell. He can no longer speak or understand English, reverting to his native language, and is generally unresponsive to basic instructions or questions. The jail moves him to a hospital to undergo psychiatric evaluation. While at the hospital, the

[1] I owe a great deal to Professor and Associate Dean Academic J. Richard Broughton for agreeing to supervise this project and for his invaluable insights and support throughout. I am also grateful to the editors of the *Lincoln Memorial University Law Review* for their care and thoughtfulness in preparing this piece for publication in their journal.

attending psychiatrist initiates treatment for what she thinks is schizophrenia and, after some time, determines the treatment was successful. The defendant is subsequently released on bond under the condition he continues to see the psychiatrist weekly and avoid contact with his wife. Immediately upon his release, he returns home with a gun and shoots his wife, killing her instantly. When police arrest him at the scene, he returns to jail and again appears to suffer a psychotic break. He smears feces and loses the ability to understand and communicate in English. He is again referred for psychiatric evaluation and, with an interpreter present, communicates a complete lack of awareness of where he is and why he is there. After administering a battery of psychometric tests, the attending psychiatrist concludes this man likely lacked cognitive capacity and impulse control at the time of the murder. The defense counsel moves for a formal plea of not guilty by insanity.

The defense and the prosecution will consult with expert witnesses who will administer psychological tests to determine the defendant's mental state at the time of the crime. Both sides present conflicting testimony, leaving it up to the fact finder to determine which side is correct. Furthermore, each expert witness uses her own choice of combination of tests in assessing the defendant, thus reducing the measure's reliability. If the defendant is indeed criminally insane, he should be found not guilty and given the help he needs. On the other hand, if he lies to avoid prison, he should be held accountable for his crime. Thus, the implications of his insanity plea are enormous, yet there is no clear consensus as to how he can establish it or what it even means.

Nevertheless, what if there was a way to peer inside the brain of a person who claims there is something so fundamentally wrong with the way he thinks that he could not form the requisite mental state at the time of the crime that would make him guilty of it? Psychological testing allows the trained clinician to determine specific personality characteristics or psychopathologies – trends of thinking and being in the world – and correlate the test's results with a past mental state.[2] However, as we shall see, many of the tests lack reliability because the administering clinician can select a

[2] Tess M. S. Neal et al., *Psychological Assessments in Legal Contexts: Are Courts Keeping "Junk Science" Out of the Courtroom?*, 20 PSYCH. SCI. PUB. INT. 135 (2019).

combination of tests to administer. Furthermore, the defendant can lie, which confounds the results. In what follows, I shall argue that cutting-edge neuroimaging and brain network analysis may provide an objective measure of a defendant's state of mind that can either correlate with the psychometric tests or stand on its own. In other words, neurotechnology can help the justice system generate a brain print – a measurable pattern of brain activity that provides insight into a person's ability to form reasonable, societally acceptable goals, act on those goals, and understand the consequences of his or her actions.

Part I discusses the two necessary components of criminal law: *mens rea* and *actus reus*. In Part II, I argue the law is correct in dichotomizing voluntary and involuntary action in the *actus reus*. This differentiation is accurate according to our current understanding of how the nervous system forms goal-directed action. In Part III, I look into *mens rea*, discussing both the Common Law and Model Penal Code categories, and argue the mental state of a "conscious object" arises in a different part of the brain than voluntary actions. Part IV looks at what happens when diseased, as seen in criminal insanity, influences *mens rea*. I discuss the history of the insanity defense and how courts assess it. Lastly, Part V lays out the current neuroscientific theories of executive function and planning. I discuss how those theories employ complex mathematical modeling that allows for future predictions of the brain network's state based on past states and how neuroimaging modalities can image brain network activity. I conclude by arguing that, using resting-state fMRI, it is possible to objectively measure *mens rea* to accurately assess a case of criminal insanity.

II. CRIMINAL ACTION: THE ACTUS REUS REQUIREMENT

A thought does not render a person criminally liable, no matter how depraved. Instead, an individual must act on that thought to commit a crime. Indeed, the law predicates criminal liability on whether a defendant engaged in a voluntary act or failed to act.[3] In this section, I discuss the *actus reus* component of a crime in the context of cognitive neuroscience.

[3] MODEL PENAL CODE § 2.01 (AM. L. INST. 1985).

I begin with the Model Penal Code (MPC), which defines an "act" or "action" in general as "a bodily movement[,] whether voluntary or involuntary."[4] However, action alone does not make a criminal. The drafters of the MPC understood this and so, in their official comments, provided a distinction between action in general and voluntary action.[5] They also note that "[the] law cannot hope to deter involuntary movement or stimulate action that cannot physically be performed."[6] Thus, an involuntary movement can justify court-mandated therapy or liability to protect the public, but an action committed involuntarily does not make a defendant criminally liable.[7] This "requirement of voluntariness, much like the mens rea limitation, is designed to limit criminal punishment only to those offenders who are truly responsible."[8] Importantly, these MPC distinctions are in agreement with several Supreme Court decisions.[9]

So, what is the difference between voluntary and involuntary action, and why does this matter to my analysis? Deborah Denno notes that voluntary actions have three defining features: (1) an internal volition; (2) an external, physical manifestation of the volition; and (3) the internal and external components are causally connected.[10] Denno implies a causal nexus between the mind and the body – a notion Denno discusses when she argues that, for an act to be voluntary, there must be a correlation between the internal and external states involved in that action.[11] For one to be criminally liable, there must be some output from the nervous system that results in the body's movement; that is to say, a nervous system output that leads to some displacement of the body or one of its members through space. Thus, criminal law carefully delineates between voluntary and involuntary acts. If an "act"

---

[4] MODEL PENAL CODE § 1.13(2) (AM. L. INST. 1985).
[5] *Id.*
[6] Ian P. Farrell & Justin F. Marceau, *Taking Voluntariness Seriously*, 54 B.C. L. REV. 1545, 1554 (2013).
[7] MODEL PENAL CODE, *supra* note 3, at § 2.01(1).
[8] Farrell & Marceau, *supra* note 6, at 1554.
[9] *See generally* Robinson v. Cal., 370 U.S. 660 (1962) (noting that differentiating between a voluntary and involuntary act is not always easy); Powell v. State of Tex., 392 U.S. 514 (1968).
[10] Deborah W. Denno, *Crime and Consciousness: Science and Involuntary Acts,* 87 MINN. L. REV. 269, 275-76 (2002).
[11] *Id.*

or "action" is the umbrella category of bodily movement, criminal liability can only be assigned when the act is *voluntary*.[12] Furthermore, voluntariness implies that the action was willed and under the individual's control and execution.

Yet volition to act is not a singular process but rather one that emerges from underlying states of the mind and the brain as a form of intentionality that takes propositions as its objects.[13] Actions are voluntary because the full range of potential movements of the body – that exist in the mind as only potentials – lead to an output that emerges from a specific brain state *chosen* by the individual.[14] For example, my arm can perform a range of movements based on its structure, size, dimensions, underlying musculature, and innervation. This range of possible movement is encoded as "states" or "codes" within my brain. When I move my arm to reach for my pen, I select one movement plan from among many and execute that plan so I may reach out and grasp my pen. There is an internal selection among possible outcomes, such that the external actualization of that outcome is causally connected to the internal state from which it arose. To my mind, this satisfies Denno's third prong for voluntary action: that there be a causal nexus between internal and external states.[15]

I contend that our current understanding of neuroscience and psychology corroborates the legal understanding of voluntary action. Thus, Hirstein and Sifferd point out that action is voluntary only insofar as it is minimally connected to the defendant's goal states or desires.[16] Another author, Susanna Blumenthal, notes the psychological capacity for voluntary action must be the product of the actor's conscious decision – it is insufficient for rationality to be read into the

---

[12] *See* MODEL PENAL CODE, *supra* note 3.

[13] *See* MICHAEL S. MOORE, ACT AND CRIME 123–124 (1993).

[14] *Id.* at 350; *see also* H.L.A. HART, PUNISHMENT AND RESPONSIBILITY: ESSAYS IN THE PHILOSOPHY OF LAW (1968) (noting an unconscious act or reflex is not ascribable as a human action since there must be a minimum nexus between the mind in the body).

[15] Denno, *supra* note 10, at 276.

[16] William Hirstein & Katrina Sifferd, *The Legal Self: Executive Processes and Legal Theory*, 20 CONSCIOUSNESS & COGNITION 156, 171 (2011).

action.[17] However, the legal standard is higher in that action must be *chosen* from several alternatives.[18]

Furthermore, there is a difference between awareness of the action occurring and one's sense of agency in its execution. For example, I may perceive my arm extending to grasp my pen and notice a difference in the position of the pen and its presence in my hand. I know something has happened as there is a change enacted in the physical world, and indeed the movement of my arm is one of many possible outcomes it can perform. However, the physical change to the world due to my bodily movement is insufficient to make *me* the actor if I only feel my arm moving. To attribute volition to action, I must realize that I am the one *causing* the action, that it is self-generated and therefore different from those caused by others.[19] Thus, the cognitive aspects of voluntary action aim toward an effect.

There is a difference between an action's effect and its intent. For Michela Balconi, an *intentional* action is a realization of an internal plan from the idea of "intentional directedness and content."[20] The concept of *voluntariness* refers to the directness or definitiveness of the action's end goal rather than forming an internal mental plan. A simple body movement, or spasm, is different from action in that the actor is in a specific relationship with his bodily movements during the time he is

---

[17] Susanna L. Blumenthal, *The Default Legal Person*, 54 UCLA L. REV. 1135, 1266 (2007).

[18] *Id.*

[19] MICHELA BALCONI, NEUROPSYCHOLOGY OF THE SENSE OF AGENCY 4 (2010). There are several neuropsychological disorders in which there is a divorce between awareness of the action and awareness of one's agency. Patients often experience a confusing world in which their limbs move independently of their agency, or they may not realize they are the ones acting. These breakdowns in pathological cognitive function give evidentiary weight to the cognitive psychological model proposed by Balconi, in which voluntary action involves the seamless connection between awareness of the act and awareness of one's agency as the actor.

[20] *Id.* Thus, the law understands intent as the "conscious object" of the action or question, which involves notions such as goals and motives. More on this in the following section on mens rea. In referring to "content," Balconi means the content of thought. Thus, intentional action follows from the internal thought to act in a certain way. This is distinguishable from volition, in which the action itself has a *telos* or end goal rather than the actor.

acting, and the core of that relation is guidance.[21] Thus, the actor engaged in voluntary action is consciously aware of what he is doing because he is guiding and controlling the action. Therefore, awareness is based on control and not on mere observation.[22] Awareness is significant, as it speaks to the actor's agency, independent of his intention. Thus, voluntary physical actions are a "sense of oneself as a physical agent producing physical effects in the world via bodily interactions with it."[23]

Modern neuroscience corroborates these psychological theories of voluntariness. The human brain's cerebral cortex contains two distinct yet interconnected regions that lie adjacent to each other and directly control movement.[24] One region, the secondary motor cortex, contains all the motor plans we have developed over our lifetime.[25] There are plans for each bodily movement for how much tension each muscle must have, its orientation in space, and the angle the joint pulls.[26] Reaching for my pen is very different from reaching for a hot cup of coffee, yet it is the same arm and hand. Walking is different from climbing stairs, yet it is the same legs that do both. Over a lifetime, particularly during early development, when the body learns how to move, the plans for movement are written or encoded into the secondary motor cortex.[27] Nevertheless, the secondary motor cortex usually does not act on its own. Instead, it must receive input from the primary motor cortex, from higher brain centers, telling it which motor

---

[21] *Id.*

[22] *Id.*

[23] *Id.*

[24] *See* Nathalie Picard & Peter Strick, *Motor Areas of the Medial Wall: A Review of Their Location and Functional Activation*, 6 CEREBRAL CORTEX 342 (1996).

[25] *Id.*

[26] *See* Patrick Haggard, *Human Volition: Towards a Neuroscience of Will*, 9 NATURE 943, 936 (2008) ("The human and primate brain both contain several distinct motor circuits that contribute to voluntary action. These circuits converge on the primary motor cortex (M1), which executes motor commands by transmitting them to the spinal cord and muscles. M1 was therefore considered to be a 'final common path' for voluntary action. It receives two broad classes of inputs, which subserve voluntary and stimulus driven actions, respectively. More recent work suggests that other cortical motor areas also send outputs to the spinal cord.").

[27] *See* Picard & Strick, *supra* note 24.

plan to select for a given task.[28] To use the pen example from above: the decision is made high in the brain's hierarchy that I need to use my pen, and the brain communicates that decision to the secondary motor cortex, which runs through the proverbial Rolodex of movement plans, carefully selecting the best one that will accomplish the task. Once the secondary motor cortex selects the plan, it communicates to lower brain structures that fine-tune it, and then it goes to the muscles, which are instructed on how to move. Thus, neuroscience confirms that for an action to be voluntary, it must be willed. In the next section, I will show how volition as the *telos* of action differs from its intent.

The legal understanding of voluntariness in a crime's *actus reus* reflects how the nervous system and human mind give rise to one's sense of agency over his or her bodily movements. The law is correct in differentiating between voluntary and involuntary action. Moreover, the notion of voluntariness is grounded in the nervous system's operation. Thus, measurable changes to nervous system functioning may result in observable behavioral differences. This causation allows me to lay the foundation for my argument that neuroimaging may provide an objective measure of criminal insanity. In what follows, I discuss the mental element of a crime.

III. MENS REA: THE GUILTY MIND

The voluntary act is not, on its own, a sufficient condition of a defendant's criminal liability. As we have seen, acts are voluntary if the actor is *aware* that he is the one acting. This awareness arises out of the neural pathways involved in a general action. [29] Psychology and neuroscience provide a convincing argument that this convergence of the awareness between the act and one's agency in committing it points to a

---

[28] *Id.*

[29] *See* People v. Grant, 377 N.E.2d 4, 8 (Ill. 1978) ("Certain involuntary acts, i.e. those committed during a state of automatism, occur as bodily movements which are not controlled by the conscious mind. A person in a state of automatism lacks the volition to control or prevent the involuntary acts," and therefore cannot be criminally responsible for such involuntary acts.); Corder v. Commonwealth, 278 S.W.2d 77 (Ky. 1955); People v. Carlo, 46 A.D. 2d 764 (N.Y. App. Div. 1974).

higher level of cognition that we may call the "will." Nonetheless, a voluntary act is insufficient to make a defendant criminally liable without an accompanying, heightened mental state.[30] In other words, acting towards an elucidated end is not criminal unless the end causes social harm, *and* it was one's goal to accomplish that harmful end. Therefore, awareness of one's actions and agency is not enough; the act must cause a consciously pre-determined, harmful end. This need for a higher mental state gives the second necessary component of a crime, the *mens rea*.

At common law, a defendant's criminal liability is asserted in the famous Latin expression *actus reus non facit reum nisi mens sit rea* – the guilty act does not make one liable unless the mind is also guilty.[31] "'There can be no crime[,] large or small, without an evil mind. It is, therefore[,] a principle of our legal system, as probably it is of every other, that the essence of an offense is the wrongful intent, without which it cannot exist.'"[32] The Supreme Court has consistently found that *mens rea* is a necessary component of criminal liability. Justice Jackson expounded in *Morissette v. United States*:

> The contention that an injury can amount to a crime only when inflicted by intention is no provincial or transient notion. It is as universal and persistent in mature systems of law as belief in freedom of the human will and a consequent ability and duty of the normal individual to choose between good and evil.[33]

I begin by tracing the origin of the common law mental requirement. Historically the early English common law of crime developed from blood feuds and was based on an innate

---

[30] *See* Farrell & Marceau, *supra* note 6, at 1584 n.208.

[31] My translation; *see also An Act Does Not Constitute a Person Guilty Unless Done with a Guilty Mind*, LAW. & JURISTS: BARRISTER, ADVOC. & LEGAL CONSULTANTS, https://www.lawyersnjurists.com/article/an-act-does-not-constitute-a-person-guilty-unless-done-with-a-guilty-mind/#:~:text=In%20Britain%20and%20other%20common,do%2C%20or%20cause%2C%20a%20criminal (last visited Oct. 7, 2022).

[32] Eugene J. Chesney, *The Concept of Mens Rea in the Criminal Law*, 29 J. CRIM. L. & CRIMINOLOGY 627 (1939).

[33] 342 U.S. 246, 250 (1952).

desire to avenge a wrongful death.[34] Indeed, Chesney observes, "[i]t is worthy of note that the criminal law concerned itself with those injuries which were highly provocative and the most injurious of these are the intentional ones."[35]

While born out of blood feuds between powerful families, the English common law, based on prior Roman Law, distinguished between intentional and unintentional crimes and ascribed higher criminal culpability to intentional ones.[36] While evident on its face, the state of the perpetrator's mind played a critical role in the common law tradition in determining the extent of their culpability and the severity of their punishment. Justice Oliver Homes notes:

> Vengeance imports a feeling of blame and an opinion, however distorted by passion, that a wrong has been done. It can hardly go very far beyond the case of a harm intentionally inflicted; even a dog distinguishes between being stumbled over and being kicked … The early English appeals for personal violence seems to have been confined to intentional wrongs.[37]

Francis Sayre adds:

> The clearest indication of criminal liability imposed by the early law without blameworthy intent is perhaps to be found in the cases of killing through misadventure and in self-defense. In early times, with the exception of killings under the king's warrant or in the pursuit of justice, which had always been justifiable, so far as we know the killer seems to

---

[34] *Id; see also* J. W. C. Turner, *The Mental Element in Crimes at Common Law*, 6 CAMB. L. J. 31, 34 (1936) ("Legal historians tell us that in the earliest periods of our law the mental state of the wrongdoer was little, if at all, regarded, and that no mental element was required to establish his liability."); *Contra* OLIVER WENDEL HOLMES, THE COMMON LAW 4 (1881).

[35] Chesney, *supra* note 32.

[36] *Id.*

[37] HOLMES, *supra* note 34, at 3.

> have been held liable for every death which he caused, whether intentionally or accidentally.[38]

Sayre further argues that up until the twelfth century, our current understanding of *mens rea* did not exist.[39] Nonetheless, he points out the consideration of the *mens rea* component.[40] He continues, "it is significant that from the earliest times of which we have any record[,] the felony of arson depended upon proof of an intent to burn."[41]

The seeds of the *mens rea* component in English law likely go as far back as the laws of Alfred, the code of laws compiled by Alfred the Great in 892 AD.[42] The code allows for an unintentional killing; "'if a man has a spear over his shoulder, and any man stake himself upon it, that he pay the *wer* without the *wite*. If he be accused of willfulness in the deed, let him clear himself according to the *wite*; and with that[,] let the *wite* abate."[43]

Indeed, this notion of criminal willfulness flourished well into the 13th century. Sayre incorporates ideas from Henrici de. Bracton's commentaries on the Common Law:

> [w]e must consider with what mind (*animo*) or with what intent (*voluntate*) a thing is done, in fact or in judgement, in order that it may be determined accordingly what action should follow and what punishment. For take away the will makes every act indifferent, because your state of mind gives meaning to your act, and a crime is not committed unless the intent to injure (*nocendi voluntas*) intervene, nor is a theft committed except with intent to steal.[44]

---

[38] Francis B. Sayre, *Mens Rea*, 45 HARV. L. REV. 974, 979-80 (1932).
[39] *Id.* at 977.
[40] *Id.* at 978.
[41] *Id.* at 981.
[42] *Id.* at 982.
[43] *Id.*
[44] *Id.* at 985. Although Bracton uses the term "will," we shall see this means something different than willful action. As previously discussed, a willed, or voluntary action, is only willed insofar as it is under the action's direct cognitive and neurological control. The actor recognizes the seamless connection between the action and its agency.

Commenting on these historical origins, Jens Ohlin notes, "[t]hese phrases link culpability with the inner blameworthiness for punishment. In order to make punishment morally justified, the actor's mind must have displayed a culpable mental state while performing the action."[45] Mental states also play an essential role in criminal law: "they provide a scheme for distinguishing between levels of culpability."[46]

Having seen how the law developed historically to add a heightened mental state requirement to criminal action, I now turn to the common law codification of this criminal mental state. At common law, the highest mental state which results in the highest criminal culpability is the nebulous and imprecise term, *malice*.[47] The word goes beyond implying an ill will against another person and instead conveys a wrongful act done intentionally with a "wicked" motive or "evil" intent.[48] The Court in *Bromage v. Prosser* understood that to act with malice meant the defendant was consciously acting in violation of the law or social norms or incited his mind to act in such a way as to harm another person.[49] The court in *Harris v. State* gives us a more precise definition in which malice "'is a condition of the mind which shows a heart regardless of social

---

There is no moral quality ascribable to voluntary action. We ascribe a moral component to an actor when they commit a voluntary action towards the end of actualizing potential harm. This is the difference between a *willed* action and an *intentional* one.

[45] JENS OHLIN, CRIMINAL LAW: DOCTRINE, APPLICATION, AND PRACTICE 142 (2d ed. 2018).

[46] *Id.*

[47] *Id.*

[48] *Id.* at 143. *See* Bromage v. Prosser, 107 Eng. Rep. 1051, 1054 (K.B. 1825), in which the court found the debtor had acted with malice when he "committed a wrongful act, done intentionally, without just cause or excuse." *See also* Regina v. Cunningham, 2 Q.B. 396 (1957), in which a defendant was arrested for tearing a gas meter from the basement of a house and stealing it, causing a gas leak that made a neighbor ill. At trial, the judge instructed the jury to infer malice as something wicked, something the defendant "has no business to do and perfectly well knows it. 'Wicked' is as good a definition as any other which you would get." Note, this instruction was found to have been made in error at appeal. The appellate court found the defendant acted with malice only if "he foresaw that the removal of the gas meter might cause injury to someone but nevertheless removed it."

[49] 107 Eng. Rep. 1051, 1054 (K.B. 1825).

duty and fatally bent on mischief, the existence of which is inferred from acts committed or words spoken.'"[50] Thus, malice in Common Law applies to an act done *wrongfully* and *intentionally* without justification.[51] An act can be willful without being malicious, but one cannot act maliciously without acting voluntarily. Building on the discussion on voluntary action, for example, reaching for a cup of coffee is a voluntary action but one without malice. When the end goal of the action is to accomplish some pre-determined, nefarious end, then the voluntary action is malicious. Thus, malice in Common Law necessarily includes the convergence of *intent* and *will* as distinct legal and neurocognitive categories.[52]

As we have seen, in common law, malice requires both *intent* and *will*; yet, applying the concept of intent has been historically thorny.[53] Unquestionably, evil intent is a staple of the common law tradition. How do we define and determine the presence of evil intent? The problem, as Ohlin points out, is that "the language of intent is fundamentally ambiguous because it is often unclear whether we are talking about intended actions or intended consequences."[54] If we understand the common law notion of *intent* as referring to intended actions, then the notion merges with the concept of willful action. Recall that voluntary action differs from involuntary action precisely because it is willed – the action is guided and controlled. In this case, we risk equivocating the word "intent" with the word "will," which we have already determined

---

[50] 8 Tex. Ct. App. 90, 109-10 (Tex. Crim. App. 1880).

[51] *See* Dunn v. Hall, 1 Ind. 371 (Ind. 1849).

[52] *See* Shannon v. Jones, 13 S.W. 477, 478 (Tex. 1890) ("'Malice means wickedness of purpose, or a spiteful or malevolent design against another, a purpose to injure another, a design of doing mischief, or any evil design or inclination to do a bad thing, or a reckless disregard of the rights of others, or an intent to do an injury to another, or absence of legal excuse, or any other motive than that of bringing a party to justice.'"); *see also* Williams v. Williams, 37 P. 614 (Colo. 1894). While these cases date to the 19th century in the U.S., they are part of a much broader Common Law heritage which we have only touched upon above. *See* Sayre, *supra* note 38, at 993 ("By the second half of the seventeenth century, it was universally accepted law that an evil intent was necessary for felony as the act itself.").

[53] *See* Chesney, *supra* note 32, at 634 ("It is safe to assert at the outset, that the general concept of mens rea necessary for criminality was very vague.").

[54] OHLIN, *supra* note 45, at 143.

means two different things. Perhaps we intuitively understand that someone who acts intentionally to cause harm ought to be punished more severely than one who acts unintentionally, but can a person act voluntarily without acting intentionally?

An example may help to illustrate the importance of avoiding a category error and conflating "intent" with "volition." If I reach for a knife, my arm extends from my body towards the object; I grasp it in my hand and lift it. This seamless integration of the action with my conscious awareness of myself as the actor is what defines voluntariness. If I reach for the knife and instead, my leg jumps out, the action uncouples from my agency, and thus the leg kick becomes involuntary or unwilled. The *telos* or *end goal* of the voluntary action is to pick up the knife. It, therefore, remains morally neutral. Action, personal agency, and the action's end goal seamlessly integrate into volition, which remains separate from the *purpose* or *intent* that motivated the action. Thus, if I pick up the knife to cut an onion, the *intent* behind the action is to cut vegetables. If I am picking up the knife to stab someone, I intend to kill. In either case, the action remains morally neutral – extending my arm and grasping the knife – despite its voluntariness; the action takes on a moral quality when seen through the lens of its *raison d'etre* – the *why* behind the action. Picking up a knife to cut an onion and then accidentally slicing my co-worker's finger is very different from picking up the knife with the desire to slice the finger. In both cases, the end goal of the *action* is to pick up the knife, yet the *purpose* that drives the action is different. One can perform a voluntary action without malice, but one cannot perform a malicious act involuntarily. The goal behind the action that spurred its execution is what determines the culpability of the actor and not whether the action was facially voluntary or not.

So, in looking at the word "malice," part of the problem is the imprecise language used in the common law to describe *mens rea*, such that terms that ought to remain distinct conflate into each other arbitrarily. While it is a historical reality that common law legal systems intuitively understood the difference in culpability depending on the defendant's mental state when he or she committed the crime, the language of "wickedness" and "malice" left too much room for judicial interpretation. So, with the inception of the Model Penal Code in 1962, the drafters' goal was to "harmonize the confusing landscape of mental states with a consistent terminology across

different jurisdictions, and to get beyond the concept of malice in favor of more precise categories."[55] The fundamental classification of mental states in the Model Penal Code is (1) acting purposely, (2) acting knowingly, (3) acting recklessly, and (4) acting negligently.[56]

The Model Penal Code ascribes acting purposely with intent, and § 2.02 defines that.[57] A person acts purposely if: (i) the element involves the nature of his conduct or a result thereof, it is his conscious object to engage in conduct of that nature or to cause such a result; and (ii) if the element involves the attendant circumstances, he is aware of the existence of such circumstances or believes or hopes they exist.[58] Thus, to act purposely is to have the conscious object of producing the result; furthermore, having a conscious object is the highest level of criminal culpability according to the Model Penal Code.[59]

IV. PURPOSE AND CONSCIOUS OBJECT – PSYCHOLOGY AND NEUROSCIENCE OF THE LAW

In what follows, I must tease apart what "conscious object" means and how cognitive psychology and neuroscience agree with the legal understanding. The addition of the term "conscious object" is the Model Penal Code's attempt to give more clarification than the common law; however, it remains an elusive term. For Ohlin, conscious object means "[t]he defendant desires to commit the act or produces the relevant result…."[60]

To say a person desires to commit an act means he or she is motivated and directs her goal toward that action. Goals, motives, and awareness of the result all converge in the notion of conscious object and direct the action to commit it voluntarily. Officially, a goal is "a desired future state (an end) coupled with a set of antecedent acts that promote the

[55] *Id.* at 142.
[56] MODEL PENAL CODE § 2.02 (AM. L. INST. 1985).
[57] *Id.*
[58] OHLIN, *supra* note 45, at 144.
[59] *Id.*
[60] *Id.*

attainment of that end state."[61] As human beings, it is intuitive to us that a conscious choice and set of internal processes produce our behaviors.[62]

In cognitive psychology, a person who acts with agency and conscious intent purposely forms judgments and decides, after careful, rational reflection, what the best course of action might be, and then initiates and guides his subsequent behavior such that it unfolds along those pre-determined, intended trajectories.[63] Thus, the antecedent acts inherent in goal-directed behavior unfold according to the actor's conscious intent and decision, desiring the action's outcome to effectuate his goal.[64] Once the actor decides on a goal, the mind must communicate the goal to other parts of the brain that initiate the sequence of events leading to the voluntary action. Conscious object, volition, and action, while intimately related, remain distinct processes but work together in goal-directed behavior.

The structure of the nervous system corroborates this psychological theory. Goal-directed behavior arises in higher areas of the cerebral cortex, which communicate continuously through areas of the brain that integrate incoming sensory information.[65] These integration areas bring together visual, auditory, somatosensory, and other senses and stitch the sense

---

[61] Elliot T. Berkman, *The Neuroscience of Goals and Behavior Change*, 70 CONSULTING PSYCH. J.: PRAC. & RSCH. 28, 44 (2018).

[62] *See* Julie Y. Huang & John A. Bargh, *The Selfish Goal: Autonomously Operating Motivational Structures as the Proximate Cause of Human Judgment and Behavior*, 37 BEHAV. & BRAIN SCIS. 121, 135 (2014).

[63] This theory of volition goes at least as far back as Sigmund Freud. *See* Roy F. Baumeister et al., *Ego Depletion: Is the Active Self a Limited Resource?*, 74 J. PERSONALITY & SOC. PSYCH. 1252, 1265 (1998) ("[Freud] described the ego as the part of the psyche that must deal with the reality of the external world by mediating between conflicting inner and outer pressures.").

[64] Even if the antecedent acts are not themselves within the conscious perception or control of the actor, their program is initiated by the actor's goal. More on this will follow in an example.

[65] *See* Nu Li & Thomas D. Mrsic-Flogel, *Cortico-Cerebellar Interactions During Goal-Directed Behavior*, 65 CURRENT OP. NEUROBIOLOGY 27, 37 (2020); *see also* Michael N. Economo et al., *Distinct Descending Motor Cortex Pathways and their Roles in Movement*, 563 NATURE 79 (2018) ("Motor cortex activity anticipates specific future movements, often seconds before the onset of movement."); Edward Evarts & Christoph Fromm, *Sensory Responses in Motor Cortex Neurons During Precise Motor Control*, 5 NEUROSCIENCE LETTERS 267, 272 (1977).

modalities into a continuously running whole, much like a movie.[66] However, this integration of the brain only plays the movie of the outside world but has no part in one's agency in that world. The integrated senses must communicate with areas of the prefrontal cortex so that goals, motivations, beliefs, and memories can integrate with the senses and an action undertaken.[67] Indeed, studies in neurobiology show goal-directed behavior rests on an individual's acquisition of two types of cognitive associations: (1) she has to learn there is a correlation between action and outcome so she may choose an action when the particular outcome becomes desirable;[68] and (2)

---

[66] *See* ANTONIO ROSSINI & CHRISTOS STRUBAKOS, REPEATING WORDS, RETELLING STORIES: REPETITION, VARIATION, AND SERIAL SIGNIFICANCE IN LITERATURE 5-19 (2021). This sensory integration takes place in what is known as the Parieto-Temporal-Occipital integration area (PTO). In neuroscience parlance, this is known as the "tertiary" cortex. As sensory information enters the nervous system through the sense organs, it undergoes higher and higher levels of processing as it rises, quite literally, higher into the brain. For example, in looking at an object in front of us, light reflecting off the object enters the eye and sends nerve impulses into the brain. These nerve impulses make their first connections in a relay station embedded deep in the brain, sending the information to the part of the brain that processes incoming visual stimuli. Each feature of a visual phenomenon gets deconstructed and processed throughout the visual cortex. Color, shape, line orientation, and sheen, all have different groups of cells that process that particular piece of information about the object in our visual field. As the relay of information gets higher and higher in the brain, the information starts to get reconstructed, or put back together, so that the object becomes apparent in our mind's eye. All sensory modalities get processed in similar ways. As the processing goes higher in the brain and becomes more complex, the senses converge in a single part of the brain, the PTO, that puts them all together and allows me to experience all sensory aspects of my current reality simultaneously. I am typing on my keyboard, I feel the keys beneath my fingertips, I hear the music playing through my headphones, and I see the screen.

[67] *Id.*

[68] *See* Bernard Balleine & Anthony Dickinson, *Goal-Directed Instrumental Action: Contingency and Incentive Learning and Their Cortical Substrates*, 37 NEUROPHARMACOLOGY 407 (1998); Henry Yin & Barbara Knowlton, *The Role of the Basal Ganglia in Habit Formation,* 7 NAT. REV. NEUROSCIENCE 464 (2006).

"the association between the outcome[] and their current motivational value."[69]

It may be helpful to explain with another example. If I feel like taking a sip of my tea, I have to weigh that desire against all other possible courses of behavior at this present moment. For example, I could take a sip of tea, make a sandwich, and move to Istanbul; there are many courses of action I *could* take, but I *decide* that it is a tea I want. This decision arises in the prefrontal cortex of my brain, in an area very close to my forehead above my eyes. Once I decide to take a sip of tea, I have to perceive where the tea is located. The tea is not a part of my body but exists apart from my limits and what I can control directly. So, I must use my body in a way that will bring the tea under my control. To do this, the part of my brain that has allowed me to decide I want tea must communicate with the sensory integration areas. There are many objects in front of me, not just tea, and I must know where each object is relative to me.

Multiple sensory modalities must be integrated simultaneously to know where objects are. Regions of my brain that assist the prefrontal cortex in carrying out the desired action continuously share all this information. Once the brain recognizes the teacup, the prefrontal cortex sends a message to the parts of the brain that control the motor system. Intent now communicates with volition. The intent, or purpose, to take a sip of tea gives an order to the part of the brain that initiates action and ensures the voluntariness of the action committed. Voluntary action then picks from among all of the movement plans encoded in the nervous system and selects the best one for reaching and grasping a cup.[70] There are a nearly infinite number of arm movements, but not all are appropriate for picking up a cup. The primary motor cortex, the part of the brain responsible for voluntary action, selects the best

---

[69] *See* Francesco Mannella et al., *The Nucleus Accumbens as a Nexus Between Values and Goals in Goal-Directed Behavior: A Review and a New Hypothesis*, 7 FRONTIERS BEHAV. NEUROSCIENCE 135, 136 (2013); *see also* Bernard Balleine et al., *The Effect of Lesions of the Basolateral Amygdala on Instrumental Conditioning*, 23 J. NEUROSCIENCE 666 (2003); Bernard Balleine & Sean Ostlund, *Still at the Choice-Point: Action Selection and Initiation in Instrumental Conditioning*, 1104 ANNALS N.Y. ACAD. SCIS. 147 (2007).

[70] *See* Edward Evarts, *Relation of Pyramidal Tract Activity to Force Exerted During Voluntary Movement*, 31 J. NEUROPHYSIOLOGY 14 (1968).

movement plan for the goal from among the Rolodex of movement plans, then gives the order to lower brain structures to communicate to the muscles involved. My arm then extends, continuously feeding back information to the brain about the arm's position in space. It then grasps the cup and raises it to my mouth so I can take a sip. A simple action that takes fractions of a second to complete involves multiple steps; each has to follow sequentially and flawlessly from the previous one to accomplish the desired end.[71]

What if I extend my arm, intending to take a sip from my cup of tea, but I knock the cup over, spilling it all over my desk? We can apply an intent-volition-action framework to this hypothetical. The purpose of the action is to take a sip from the cup. The action is voluntary in that the arm extends, yet the action's end does not match its intended goal because the *result* was unintended.

Thus, it is worth noting from the above example that each time a goal-directed action occurs, specialized groups of cells in specific locations in the brain act to ensure the whole process occurs seamlessly. In other words, these locations are anatomically conserved, which means they are in the same spot of the brain in almost all people and do the same job for everyone. Moreover, they work together, forming a network of brain areas that guides goal-directed action.

---

[71] *See* Karel Svoboda & Nuo Li, *Neural Mechanisms of Movement Planning: Motor Cortex and Beyond*, 49 CURRENT OP. NEUROBIOLOGY 33 (2018) ("During perceptual decision-making, behavior-related information flows from sensory to motor areas. Decisions occur in parts of the brain where correlations between neural activity and future actions first emerge . . . . During the delay, neurons in frontal and parietal cortex and connected structures show persistent and ramping activity related to specific movement, long before movement onset. This neural correlate of future movement is referred to as 'preparatory activity' . . . ."); *see also* Krisna V. Shenoy et al., *Cortical Control of Arm Movements: A Dynamical Systems Perspective,* 36 ANN. REV. NEUROSCIENCE 337 (2013) (arguing for a dynamic system-based approach to movement initiation in which previous states of the system influence future ones).

V. PURPOSE AND CONSCIOUS OBJECT – CRIMINAL AND EVIDENTIARY FRAMEWORK

As we have seen, the notion that people have mental states from which we can ascribe degrees of culpability has a history rooted in Roman and English common law. In response to the confusion generated by the imprecise Common Law definitions, in 1962, the American Law Institute published the Model Penal Code, which modernized *mens rea* into purpose, knowledge, recklessness, or negligence.[72] Furthermore, the MPC required a *mens rea* component for each element of an offense, building on the idea that "wrongdoing must be conscious to be criminal."[73]

Supreme Court cases reiterated the MPC's importance of mens rea; one such case was *Elonis v. United States*.[74] There, "the Court incorporated the MPCs four *mens rea* standards when explaining its outcome and reasoning."[75] However, mental states remain elusive not only jurisprudentially but also evidentiarily. In other words, while courts appreciate the difference in mental states outlined by the MPC, the criminal justice system must rely on direct and indirect evidence of the defendant's mental state when the defendant committed the crime.[76] This reliance raises a fundamental issue for any fact-finder or party in a criminal proceeding: attorneys must piece together evidence to persuade a jury that an unmeasurable condition in the past did or did not exist, that is, an individual's state of mind. Years may have passed since the defendant committed the crime. Yet, because of the fundamental nature of the *mens rea* element, a prosecutor must prove the defendant's state of mind at the time the defendant committed the crime beyond a reasonable doubt.

---

[72] MODEL PENAL CODE, *supra* note 3.
[73] Morissette v. United States, 342 U.S. 246, 252 (1952).
[74] 575 U.S. 723 (2015).
[75] Deborah Denno, *Concocting Criminal Intent,* 105 GEO. L. J. 323, 332 (2017).
[76] *See* Julie Schmidt Chauvin, Comment "*For It Must Seem Their Guilt": Diluting Reasonable Doubt by Rejecting the Reasonable Hypothesis of Innocence Standard,* 53 LOY. L. REV. 217, 221–22 (2007).

Most of the evidence involving a mental state is *de facto* circumstantial.[77] While courts have allowed circumstantial evidence, giving it the same weight as direct evidence, circumstantial evidence forces the jury to make a leap of logic and infer the existence of a fact at issue, connecting a circumstantial fact to a directly incriminating fact.[78] For example, an eyewitness who sees a person walk up to a woman sitting at a restaurant and shoot her before walking away provides, through their testimony, direct evidence of the event. Indeed, in this example, one could reasonably infer the defendant murdered with a conscious object or purpose since there were a series of antecedent steps the murderer took before the action's completion. Walking up to the victim, removing the gun, aiming it at the victim, pulling the trigger, and walking away, are all acts that fall under direct evidence since it is an eyewitness reporting them. However, the fact-finder must infer the defendant's intentionality because no one can give direct evidence of the defendant's mind. Indeed, any reasonable person would have no difficulty inferring intentionality, yet the culmination of the actions considered to prove *mens rea* can only do so circumstantially.[79] When evaluating all evidence, one must make an inference, but the more the inferences pile onto each other, the further removed the fact-finder becomes, and

---

[77] DEBORAH JONES MERRITT & RIC SIMMONS, LEARNING EVIDENCE: FROM THE FEDERAL RULES TO THE COURTROOM 14 (2018) ("Circumstantial evidence is any evidence that requires the jury to make an inference connecting the evidence with a disputed fact. Direct evidence, in contrast, requires no inferential bridge; it directly establishes a contested fact."). Courts have readily found that both direct and circumstantial evidence should be given equal weight. *See* Holland v. United States, 348 U.S. 121, 140 (1954).

[78] Joshua Greene & Jonathan Cohen, *For the Law, Neuroscience Changes Nothing and Everything*, 359 PHIL. TRANSACTIONS ROYAL SOC'Y B, 1775 (2004).

[79] Courts consider circumstantial evidence as being necessary for proof of the mental state. *See* United States v. Sullivan, 522 F.3d 967, 978 (9th Cir. 2008) ("Intent may be established through circumstantial evidence."); United States v. Stoker, 706 F.3d 643, 646 (5th Cir. 2013) ("'Intent may, and generally must, be proven circumstantially.'" (*quoting* United States v. Maggitt, 784 F.2d 590, 593 (5th Cir. 1986))); United States v. Smith, 508 F.3d 861, 867 (8th Cir. 2007) ("'[A] jury rarely has direct evidence of a defendant's knowledge, [and] it is generally established through circumstantial evidence.'" (*quoting* United States v. Ojeda, 23 F.3d 1473, 1476 (8th Cir. 1994))).

the easier it is to lead them into factual distortions.[80] This scenario presents us with an obvious problem: on the one hand, intuitively, historically, and very likely mechanistically, the law established varying mental states that are necessary conditions for criminal liability and which determine culpability accordingly. On the other hand, these states are necessary conditions that one cannot directly prove; instead, one must rely on a series of inferences, into any one of which an opposing party may introduce subtle distortions of the facts.[81]

Say, for example, a defendant is charged with killing his wife. The events leading up to the alleged murder were as follows: the defendant left the apartment where he was staying, entered his car, drove to his wife's home, exited the car, walked up to the house, entered the house, removed a gun from his pocket, aimed it at his wife, and pulled the trigger. Under the MPC, the prosecution would charge the defendant with first-degree murder, arguing he acted purposely to kill his wife. To get a conviction, the prosecutor must convince a jury beyond a reasonable doubt on every element of the crime. Nevertheless, how do you convince people that he had a conscious object to murder his wife when the defendant committed the crime? The prosecution must infer from the facts surrounding the action and then convey that inference to the jury, who determine culpability based on the inference. Indeed in this example, inferring intent is easy. The more steps the defendant took in preparing and executing the end goal, the easier it is to infer

---

[80] *See* Sanford H. Kadish, *The Decline of Innocence,* 26 CAMBRIDGE L. J. 273 (1968).

[81] *See* Matthew R. Ginther et al., *Decoding Guilty Minds: How Jurors Attribute Knowledge and Guilt*, 71 VAND. L. REV. 241, 245 (2018) (In this essay, the authors designed experiments to assess how well jurors could differentiate between legal mental states. The authors found "that subjects' intuitions about the level of culpability warranting criminal punishment diverge significantly from prevailing legal practice; subjects tend to regard recklessness as a sufficient basis for punishment even under circumstances where the legislatures and courts tend to require knowledge."); *see also* Jenny E. Carroll, *The Problem with Inference and Juvenile Defendants*, 45 FLA. ST. U. L. REV. 1, 4–5 (2017) ("[F]acts . . . are as much a product of the evidence used to support [the fact finder] as the inferences drawn from them. In this aspect, procedural safeguards risk failure not because they lack rigor (though they may) but because they overlook a critical component of the substantive law they apply to – the fact finder's interpretation.").

from the evidence – circumstantial though it may be – that the defendant's actions all arose from the same mental goal or plan. But what happens when the inference is wrong? What happens when all of the evidence surrounding the event points to one mental state driving the actions, yet that mental is absent or not rooted in reality? In what follows, I shall discuss the insanity defense.

VI. THE INSANITY DEFENSE – NEGATING THE GUILTY MIND?

The insanity defense is a natural outgrowth of our previous discussion on the role of agency and intent in a criminal action. Jacques M. Quen notes, at least from the time of Aristotle, that free will was critical in evaluating guilt, and free will is a mental capacity lacking in animals, children, and the insane.[82] The insanity defense's first significant milestone arose in British courts in 1724.[83] The Crown charged Edward Arnold ("crazy Ned") with attempting to murder Lord Onslow, whom he had shot and wounded.[84] In the trial transcript, Arnold claimed he was delusional and that Lord Onslow tried to kill him through witchcraft.[85] He claimed, "he plagues me day and night. I can't eat or drink; if I eat anything, it comes out of my body. I am . . . as they pumped the breath out of my body." [86] On the other hand, the prosecution argued Arnold was wicked and not insane.[87] Justice Tracy instructed the jury:

> Punishment is intended for example, and to deter other persons from wicked designs; but the punishment of a madman, a person that hath no design, can have no example . . . . [I]t is not every frantic humour or something unaccountable in a man's actions, that points him to be such a madman as is to be exempted from punishment: it must be a man that is totally deprived of his

---

[82] *See* H.B. Dearman, *Criminal Responsibility and Insanity Tests: A Psychiatrist Looks at Three Cases*, 47 VA. L. REV., 1388, 1398 (1961).
[83] Rex v. Arnold, 16 How. St. Tr. 695 (1724).
[84] *Id.*
[85] *Id.*
[86] THOMAS BAYLY HOWELL, A COMPLETE COLLECTION OF STATE TRIALS 721 (1816).
[87] *Id.*

> understanding and memory, and doth not know what he is doing, no more than an infant, than a brute, or a wild beast; such a one is never the object of punishment; therefore, I must leave to your consideration, whether the defendant . . . knew what he was doing and whether he was able to distinguish whether he was doing good or evil.[88]

Arnold was subsequently found guilty.[89] Yet the Arnold case established a test in the British Common Law system for the first time, which subsequently became known as the "Wild Beast Test."[90] To be found insane, the defendant cannot know what he is doing beyond an infant, a brute, or a wild beast who could appreciate their actions.[91] The Wild Beast Test became the "first formal legal definition of insanity."[92]

Nevertheless, the story of criminal insanity had just begun with Arnold. On a cold day in January 1843, at Saint Martin's parish in Middlesex, a man named Daniel M'Naghten took out his gun and shot Edward Drummond, whom he believed to be the Prime Minister of Britain, wounding him.[93] Drummond died several days later, the Crown formally charged M'Naghten with murder, to which he pleaded not guilty by reason of insanity.[94] During the trial, witnesses came forward attesting M'Naghten was not in his right frame of mind when he committed the act.[95] Moreover, medical evidence showed people like M'Naghten may *appear* to be of sound mind while simultaneously suffering from "morbid delusions."[96] While such a person would typically know the difference between right and wrong in all other acts, acts related directly to their delusional view of reality occur without their intent and a moral compass. In other words, during his day-to-day life, M'Naghten could appreciate the consequences of his actions

---

[88] *Id.* at 764–65.
[89] *Id.*
[90] LAWRENCE S. WRIGHTSMAN & STEPHEN PORTER, FORENSIC PSYCHOLOGY 193 (2001).
[91] *Id.*
[92] *Id.*
[93] M'Naghten's Case, 8 Eng. Rep. 200 (H.L. 1843).
[94] *Id.* at 201.
[95] *Id.*
[96] *Id.*

and ascribe morality to them regarding whether they are inherently right or wrong. However, any action relating to his delusions about the British Prime Minister lost the moral lens through which a reasonable person views all his actions and the perception of his agency in effectuating the act. Thus, the defense argued M'Naghten was not in control when he acted against the man he believed to be the Prime Minister.[97] Moreover, the defense introduced evidence before the Court that these delusions may gradually take hold in the mind and crescendo when they spill out, as it were, and commit an act against the object of the mind's delusion.[98] Lord Chief Justice Tindale noted:

> [t]he question to be determined is, whether at the time the act in question was committed, the prisoner had or had not the use of his understanding, so as to know that he was doing a wrong or wicked act. If the jurors should be of the opinion that the prisoner was not sensible, at the time he committed it, that he was violating the laws of both God and man, then he would be entitled to a verdict in his favour: but if, on the contrary, they were of opinion that when he committed the act he was in a sound state of mind, then their verdict must be against him.[99]

The Court subsequently found Daniel M'Naghten not guilty by reason of insanity.[100] Following the verdict, the House of Lords issued an advisory opinion in which they set out the legal standard for finding defendants not guilty because they are criminally insane:

> to establish a defense on the ground of insanity, it must be clearly proved that, at the time of committing 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

[97] M'Naghten's Case, 8 Eng. Rep. 200, 201 (H.L. 1843).
[98] *Id.*
[99] *Id.* at 202.
[100] *Id.*

> know it, that he did not know that what he was doing was wrong.[101]

The advisory opinion raises a couple of vital features that a defendant must meet to establish an insanity defense: (1) the defect affecting reason has to exist at the time the crime was committed; and (2) the defect must either prevent the defendant from understanding his agency in action or if he could appreciate his agency, the defect prevented the defendant from understanding that his act was wrong.[102] Ohlin summarized the M'Naghten test succinctly when he noted the test:

> requires a finding that the defendant suffered from a mental disease or defect that caused a cognitive impairment. However, there are two prongs, either one of which would be sufficient to meet the standard. Under the first prong, the defendant believes she is committing one action but is actually committing another . . . . Under the second prong, the defendant correctly perceives this action but is unaware it is wrongful.[103]

The M'Naghten test is significant because it provides a metric against which a criminal defendant's mental state may be assessed. The test provides necessary and sufficient conditions so a judge can determine the legitimacy of a defendant's *mens rea* claim. Despite its attempts at being helpful, the test has some apparent ambiguities. The first issue is that the test introduces, for the first time, the insane delusion rule.[104] For example, suppose a man kills his friend because he believes the government has sent his friend to extract his

---

[101] *Id.* at 210.

[102] OHLIN, *supra* note 45, at 811.

[103] *Id.* at 812; *see also* WRIGHTSMAN & PORTER, *supra* note 90, at 192–93 ("[T]he McNaughton (sic.) test is called a cognitive test of insanity because it emphasis the quality of the person's thought processes and perceptions of reality at the time of the crime.").

[104] *See* E. Lea Johnston & Vincent T. Leahey, *The Status and Legitimacy of M'Naghten's Insane Delusion Rule*, 54 U. C. DAVIS L. REV. 1777, 1787 (2021) ("Prior to *M'Naghten*, no court or textbook had articulated an insanity rule specific to delusions, certainly not one that subjected the delusion to an objective standard of justification or excuse.").

thoughts from his brain while acting delusional. In that case, a court may acquit the man on insanity grounds under M'Naghten. Thus, the insane delusion rule essentially functions as a mistake of fact defense. If the defendant in the above example had reasonable grounds to believe the facts were as he perceived them, he would be acting in self-defense. [105]

The second issue with the M'Naghten test centers around the term, "wrong." Ohlin considers whether this use of the term "wrong" suggests "morally wrong" or "legally wrong". Because of this ambiguity, one can understand the cognitive test in one of two ways: either the mental disease or defect prevented the actor from understanding that his actions were *morally* wrongful, or that his actions were *illegal*.[106] Indeed, this confusion was present from the onset of the case's opinion in which two judges in the House of Lords restricted the word "wrong" to "illegal" while another two used it to refer to an illegal, moral wrong.[107] While this distinction is subtle, it is nonetheless essential. Ohlin notes, "[a]t issue is whether the jury is supposed to ask whether the defendant's impairment was an obstacle to engaging in coherent moral reasoning or accurate legal analysis."[108]

These criticisms led some jurisdictions to expand the M'Naghten test beyond a test for cognitive capacity to include a volitional component.[109] This evolution results from the law correctly identifying a neurocognitive difference between *purpose* and *volition*.[110] Purpose refers to the *mens rea* and

---

[105] *Id*. *See also* ROLLIN M. PERKINS & RONALD N. BOYCE, CRIMINAL LAW 966 n.89 (3d ed., 1982) ("One not suffering from an insane delusion would not be excuse[d] for such a killing unless the mistake was a reasonable one under the circumstances, but the delusion will take the place of reasonable grounds for the belief in the mistake of fact defense.").

[106] OHLIN, *supra* note 45, at 813.

[107] *See* Johnston & Leahey *supra* note 104, at 1789 n.50 (quoting M'Naghten Case, 8 Eng. Rep. 200 (H.L. 1843) ("the defendant is 'punishable according to the nature of the crime committed, if he knew at the time of committing such crime that he was acting contrary to law; by which expression we understand your Lordship to mean the law of the land.'")).

[108] OHLIN, *supra* note 45, at 813.

[109] OHLIN, *supra* note 45.

[110] *Id*.

volition refers to the *actus reus*.[111] The expanded test became known as the "irresistible impulse test."[112] Simply put, this test states, "[i]f a defendant demonstrated cognitive knowledge of right or wrong, he or she could still be found not guilty by reason of insanity if his or her free will was so destroyed or overruled that the person had lost the power to choose between right and wrong."[113] Ohlin notes that today this test is retained in a handful of states, namely Georgia, New Mexico, and Virginia.[114]

VII. HOW IS CRIMINAL INSANITY CURRENTLY ASSESSED?

The M'Naghten test continued as the criminal insanity gold standard for a century until March 30, 1981. On that day, John Hinckley Jr. shot at President Ronald Reagan six times after Reagan gave a speech at the Washington Hilton.[115] One of the bullets ricocheted off the side of Reagan's limo and hit Reagan in the chest.[116] At Hinckley's trial in 1982, he was found not guilty by reason of insanity.[117] Although both the prosecution and defense presented evidence regarding Hinckley's mental state at the time of the crime, the defense ultimately prevailed by applying the M'Naghten test.[118] Hinckley claimed he was obsessed with the actress Jodi Foster and hoped assassinating Reagan would impress her.[119] In other words, when he committed the crime, this delusion affected Hinckley's cognitive abilities concerning any actions related to the delusion, which prevented Hinkley from understanding the

---

111 *Id*.
112 *Id*.
113 WRIGHTSMAN & PORTER, *supra* note 90.
114 OHLIN, *supra* note 45, at 812.
115 Doug Linder, *The Trial of John W. Hinckley, Jr.*, FAMOUS AM. TRIALS (2008), http://law2.umkc.edu/faculty/projects/FTrials/hinckley/hinckleyaccount.html.
116 *Id*.
117 *Id*.
118 Michael M. O'Hear, *From M'Naghten to Hinckley to Clark: "The Incredible Shrinking Insanity Defense*," MARQUETTE UNIV. L. SCH. FAC. BLOG (June 18, 2010), https://law.marquette.edu/facultyblog/2010/06/from-mnaghten-to-hinckley-to-clark-the-incredible-shrinking-insanity-defense/.
119 Linder, *supra* note 115.

wrongfulness of his action. However, this last point remains unclear in Hinkley's case. Recall from above that "wrongfulness" under M'Naghten was interpreted as either illegality alone or an illegal, moral act.[120] In his letter to Jodi Foster, Hinkley fully understands his action's illegality and moral turpitude. He writes,

> There is a definite possibility that I will be killed in my attempt to get Reagan. It is for this very reason that I am writing you this letter now . . . I will admit to you that the reason I'm going ahead with this attempt now is because I just cannot wait any longer to impress you. I've got to do something to make you understand, in no uncertain terms, that I am doing all of this for your sake! By sacrificing my freedom and possibly my life, I hope to change your mind about me.[121]

Following the trial, Hinckley was committed to a psychiatric hospital for treatment.[122]

The public outcry against the jury's verdict was both immediate and intense.[123] Consequently, Congress passed the Federal Insanity Defense Reform Act in 1984, which narrowed the standard for insanity in federal jurisdictions.[124] The Act revised the M'Naghten test by adding the requirement that "at

---

[120] In his letter to Jodi Foster, Hinkley understands both his action's illegality and moral turpitude fully. He writes, "There is a definite possibility that I will be killed in my attempt to get Reagan. It is for this very reason that I am writing you this letter now…. I will admit to you that the reason I'm going ahead with this attempt now is because I just cannot wait any longer to impress you. I've got to do something to make you understand, in no uncertain terms, that I am doing all of this for your sake! By sacrificing my freedom and possibly my life, I hope to change your mind about me." *Id.*

[121] Newsweek Staff, *John Hinckley's Last Love Letter, to Jodie Foster,* NEWSWEEK (July 27, 2016, 6:16 PM), https://www.newsweek.com/john-hinckley-love-letter-jodie-foster-reagan-assassination-484716.

[122] Linder, *supra* note 115.

[123] *Id.*

[124] *Id.*

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 of the wrongfulness of his acts. Mental disease or defect does not otherwise constitute a defense."[125] In sum, the Act replaced the "delusion" standard of M'Naghten with "mental disease or defect" that specifically affects the defendant's ability to understand the nature and quality of the wrongfulness of his actions. Note it is not sufficient under the Act for the defendant to have just any mental disease or defect. The nature of the mental disease at the time of the crime must *directly* relate to the defendant's ability to assess whether his actions are right or wrong.

While helpful, the M'Naghten test also brought a unique set of problems. For example, the word "know" is imprecise because a defendant might be able to report his action is wrong without having a deep understanding of what that means.[126] Similarly, M'Naghten emphasizes cognitive capacity such that a defendant might only minimally understand or have the ability to control their conduct.[127] The final issue with M'Naghten is that it requires, on its face, a total lack of knowledge or cognitive capacity.[128]

To create a comprehensive standard for criminal insanity, the American Law Institute (ALI) advanced a new definition of insanity that arose in *United States v. Brawner*.[129] The standard states, "[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 either to appreciate the criminality [wrongfulness] of his conduct or to conform his conduct to the requirements of the law."[130]

This principle became known as the ALI standard of insanity.[131] The ALI standard is the Model Penal Code's

---

[125] 18 U.S.C. § 17(a).

[126] Eugene M. Fahey et al., *"The Angels that Surrounded my Cradle": The History, Evolution, and Application of the Insanity Defense*, 68 BUFF. L. REV. 805, 828 (2020).

[127] *Id.*

[128] *Id.* at 828-29.

[129] 471 F.2d 969 (D.C. Cir. 1972).

[130] MODEL PENAL CODE § 4.01(1) (AM. L. INST. 1985).

[131] Cornell Law School, *Insanity Defense*, LEGAL INFO. INST., https://www.law.cornell.edu/wex/insanity_defense (last visited Nov. 16, 2021).

approach to insanity, which has been described as "a concerted effort to reconcile the various details and emphases present in the traditional tests for insanity."[132] Moreover, the MPC's language includes both a cognitive and volition-based component.[133] The cognitive component is included in the "appreciate" language of the standard, and the volitional component is included in the "conform his conduct to the requirements of the law" language of the standard.[134]

In 2006, the Supreme Court decided yet another case involving the insanity defense. Eric Clark shot and killed a local police officer during a routine traffic stop believing that aliens invaded his hometown in Arizona.[135] At his bench trial, the Court convicted Clark of first-degree murder after he unsuccessfully presented the affirmative defense of insanity.[136] As part of Clark's defense, a psychiatrist testified that Clark had paranoid schizophrenia at the time of the crime and thus lacked the specific intent to shoot and kill the officer.[137] However, the State's psychiatrist thought Clark's disorder did not prevent him from "appreciating the wrongfulness of his conduct."[138]

Ultimately, the trial court did not believe Clark's defense and found him guilty of murder, sentencing him to 25 years in prison.[139] The Supreme Court granted *certiorari* on whether the court erred in relying upon the Arizona Supreme Court's decision in *State v. Mott*[140] that did not allow evidence of any mental disorder other than insanity to negate the *mens rea* of a crime. In a 5-4 decision, the Court found that Arizona's prohibition against allowing evidence of diminished mental capacity did not violate due process.[141] The Court noted that evidence regarding mental capacity and mental disease was usually based on expert testimony rather than direct

---

132 *Id*.
133 *Id*.
134 *Id.*
135 Clark v. Arizona, 548 U.S. 735 (2006).
136 *Id*.
137 *Id.* at 742.
138 *Id*.
139 *Id*. at 746.
140 931 P.2d 1046 (Ariz. 1997).
141 Clark v. Arizona, 548 U.S. 735 (2006).

observation.[142] The Court found that expert testimony could present an evidentiary danger in that an expert cannot accurately define the precise mental state when the defendant committed the crime.[143] Furthermore, the law's definition of insanity is not the same as the diagnostic categories in psychiatry, which could result in clinicians and the justice system speaking two different languages.

The Court's line of reasoning in *Clark* follows the general inferential nature of *mens rea* evidence, as previously discussed. It is worth noting that the Court takes a more stringent line regarding the inclusion of psychiatric evidence because of the inherent impossibility of providing direct evidence of a mental state at some point in the past.[144] However, I contend that these problems and pitfalls have the potential to arise in all attempts to prove a defendant's mental state, and part of the problem is the vague statutory language. Applying the vague Model Penal Code standards is a much more difficult task than initially anticipated because there is no societally

---

[142] *Id.* at 742, 737–38 ("*Mott* held that testimony of a professional psychologist or psychiatrist about a defendant's mental incapacity owing to mental disease or defect was admissible, and could be considered, only for its bearing on an insanity defense, but could not be considered on the element of *mens rea*. Of the three categories of evidence that potentially bear on *mens rea* – (1) everyday "observation evidence" either by lay or expert witnesses of what Clark did or said, which may support the professional diagnoses of disease and, in any event, is the kind of evidence that can be relevant to show what was on Clark's mind when he fired his gun; 2) "mental-disease evidence," typically from professional psychologists or psychiatrists based on factual reports, professional observations, and tests about Clark's mental disease, with features described by the witness; and (3) "capacity evidence," typically by the same experts, about Clark's capacity for cognition and moral judgment (and ultimately also his capacity to form *mens rea*)—*Mott* imposed no restriction on considering evidence of the first sort, but applies to the latter two. Although the trial court seems to have applied the *Mott* restriction to all three categories of evidence Clark offered for the purpose of showing what he called his inability to form the required *mens rea,* his objection to *Mott's* application does not turn on the distinction between lay and expert witnesses or the kinds of testimony they were competent to present. Rather, the issue here is Clark's claim that the *Mott* rule violates due process.").

[143] *Id*. at 776.

[144] *See* FED. R. EVID. 803(3).

agreed-upon standard regarding who should be acquitted and punished.[145]Indeed, although the Model Penal Code and other tests of insanity discussed above state, "[e]vidence that the defendant suffered from a mental disease or defect is admissible whenever it is relevant to prove that the defendant did or did not have a state of mind that is an element of the offense,"[146] courts often exclude such evidence finding it irrelevant to mens rea.[147]

Legally, insanity and *mens rea* can be at odds. For example, imagine a defendant found in an apartment where the surrounding circumstances suggest he was there to steal from the victim.[148] After being charged with burglary, he argued he was operating under a delusion and believed that the apartment and the property were his.[149] If the criminal statute necessitates that burglary occurs with the intent of entering a dwelling that is not one's own to steal property that is not one's own, we can see how Wetmore's delusions would negate the mens rea of his crime. Nevertheless, herein lies a logical paradox: the delusions give him an affirmative defense to a crime where those same delusions cancel his mental state. Indeed, failure to prove the crime's *mens rea* means the prosecution has not proved the defendant committed the crime.

Does this mean that Wetmore is not criminally liable, no matter how you slice it? While called "diminished capacity," cases such as this are misleading because rather than dealing with a lesser form of criminal insanity, they refer to the lack of a requisite mental element of the crime, like a mistake of fact.[150] Instead, as Stephen Morse notes, "mental disorder produces crazy desires or crazy beliefs about reality, but it virtually never prevents a defendant from meeting the law's criteria for intention, knowledge, conscious awareness,

---

145 Steven R. Smith & Robert G. Meyer, *Law, Behavior, and Mental Health: Policy and Practice*, 4 J. CONTEMP. HEALTH L. & POL'Y 496 (1988).

146 MODEL PENAL CODE § 4.02(1) (AM. L. INST. 1985).

147 *See* Erica Beecher-Monas & Edgar Garcia-Rill, *Actus Reus, Mens Rea, and Brain Science: What Do Volition and Intent Really Mean?*,106 KY. L. J. 265 (2017).

148 People v. Wetmore, 583 P.2d 1308 (Cal. 1978).

149 *Id*. at 1310-11.

150 Stephen J. Morse, *Undiminished Confusion in Diminished Capacity*, 75 J. CRIM. L. & CRIMINOLOGY 1, 6 (1984).

and other mens rea terms."[151] For example, recall how M'Naghten believed that the Prime Minister's secretary was the Prime Minister.[152]

In more recent history, Andrea Yates was convinced that she had to kill her children; otherwise, they would become corrupted and their souls tormented by the devil.[153] In other words, she knew that they were human beings, and in fact, those human beings were her children; furthermore, she knew that by drowning them she would kill them. [154] She also knew that killing them was morally and legally wrong.[155] Yet she wrongly or, under delusion, believed that killing her children was the only option to save them from eternal damnation.[156] In this case, the intent-volition neurocognitive workflow discussed above would remain intact, and yet, to a reasonable outside observer, there is something inherently wrong with Yates' intent.[157] It is not motivated by, nor directed to, an outcome that we would consider "normal." And herein lies the problem: insanity is an affirmative defense – an excuse – for a defendant who acted with sufficient mens rea to be otherwise found guilty. However, not every mental health disorder questions the workings of mens rea to qualify for the affirmative defense of insanity, as seen in the Yates example above.[158]

I agree with Craig Stern that a "properly formulated insanity defense 'show[s] the absence of an evil mind and therefore of criminal liability.'"[159] Stern also notes, "[t]he insanity defense that is rooted in the retributive basis for the doctrine of mens rea entails a cognitive test for the defense. The defect in the defendant's understanding of what went on in the

---

[151] Stephen J. Morse & Morris B. Hoffman, *Criminal Law: The Uneasy Entente between Legal Insanity and Mens Rea: Beyond Clark v. Arizona*, 97 J. CRIM. LAW & CRIMINOLOGY 1071, 1087 (2007).
[152] M'Naghten's Case (1843) 8 Eng. Rep. 200 (H.L.).
[153] Phillip Resnick, *Andrea Yates Case: Insanity on Trial*, 55 CLEV. ST. L. REV. 147, 149 (2007); Deborah W. Denno, *Who is Andrea Yates? A Short Story About Insanity,* 10 DUKE J. GENDER L. & POL'Y 1 (2003).
[154]Morse & Hoffman, *supra* note 151, at 1090.
[155] *Id*. at 1093.
[156] *Id*.
[157] *Id*. at 1094.
[158] *See* Craig A. Stern, *The Heart of Mens Rea and the Insanity of Psychopaths,* 42 CAP. U. L. REV. 619 (2014).
[159] *Id*. at 651.

commission of the offense–a defect owing to mental illness–must be taken into account if the assessment of the defendant's mens rea is not to lead to a mistaken assessment of his will."[160] Thus, the cognitive test for insanity that Stern proposes links insanity back to the *mens rea* because, as Dressler points out, such a test can evaluate how a defendant's will was perhaps influenced by faulty cognition.[161] Therefore, if a mental disease prevented the communication of significant facts from reaching the defendant's mind, the defendant relied on faulty cognition even if the intent was present.[162] To my mind, this is a more nuanced and accurate way to approach criminal insanity. In other words, this approach emphasizes that neither intent nor will exist in isolation but rather rest on a network of emotional and cognitive experiences that ensure it is directed towards socially acceptable ends.

The question, then, is how can psychiatrists and courts evaluate this faulty cognition? In other words, how can psychiatrists assess a defendant's mental state under the legal standard of insanity, especially in situations where the legal definition of intent is present – like in the case of Andrea Yates? I begin with the basics. James L. Knoll and Phillip J. Resnick advise:

> When evaluating knowledge of wrongfulness, the psychiatrist should carefully analyze the defendant's behaviors, statements, and motives. For example, hiding evidence, lying about the offense, and fleeing from the police all suggest that the defendant knew that his behavior was legally wrong. In contrast, committing a crime with no rationale motive, making no effort to avoid detection, and making no effort to flee may suggest a lack of knowledge of wrongfulness. . . . Statements made by the defendant months later that he knew the act was wrong are helpful, but care must be used to

---

[160] *Id*. at 653.

[161] JOSHUA DRESSLER, UNDERSTANDING CRIMINAL LAW 343–44 (6th ed. 2012).

[162] *Id*.

> determine whether the defendant can accurately recall his thinking at the time of the crime.[163]

Furthermore, Knoll and Resnick argue the psychiatrist must try to determine the defendant's motive for committing the crime by comparing their stated motive with ordinary criminal motives, like revenge.[164] Indeed, the Supreme Court ruled in *Daubert v. Merrell Dow Pharmaceuticals*[165] that expert opinions must be based on "an inference or assertion derived by the scientific method," and courts must consider "whether the reasoning or methodology underlying the testimony is scientifically valid and whether that reasoning or methodology can properly be applied to the facts at issue."[166] This approach may be helpful when conducted under a trained clinician's watchful eyes. The problem, however, is the psychiatrist must still infer the intent from the surrounding circumstances and facts. Moreover, an action may support multiple inferences depending on the observer's perspective, who may unconsciously shift perspective. So, how can *mens rea* be established in the context of insanity?

Because the Court in *Daubert* allowed expert opinions based on an inference when that inference is scientifically sound, psychiatrists and psychologists now rely on increasing tests to peer into a defendant's mind and determine whether they are criminally insane.[167] The most common test administered to a criminal defendant pleading not guilty because of insanity is the Minnesota Multiphasic Personality Inventory – 2 (MMPI-2).[168] Initially developed in the 1960s as a personality assessment measure, the MMPI-2 is a psychometric test that looks at ten scales, each representing 10 major categories of abnormal behavior and psychopathology, as well as four

---

[163] James L. Knoll IV & Phillip J. Resnick, *Insanity Defense Evaluations - Basic Procedure and Best Practices,* PSYCHIATRIC TIMES (Dec. 1, 2008), https://www.psychiatrictimes.com/view/insanity-defense-evaluations-basic-procedure-and-best-practices.

[164] *Id.*

[165] 509 U.S. 579 (1993).

[166] *Id.* at 592-93.

[167] *See* Kumho Tire Co. v. Carmichael, 526 U.S. 137 (1999).

[168] *See* JAMES N. BUTCHER & CAROLYN L. WILLIAMS, ESSENTIALS OF MMPI-2 AND MMPI-A INTERPRETATION 196 (2d ed. 2000).

validity scales designed to detect lying.[169] Because of tests like the MMPI-2 that have probative value in assessing personality traits, psychological testing has been implemented in the legal setting beyond just criminal cases. Unfortunately, there are serious questions regarding many of these tests' scientific validity and reliability.[170] Scientific reliability is a measure of the test's internal consistency, which means the test is administered to multiple people or across multiple time points, rendering similar results.[171] Validity asks whether the test measures what it claims to be measuring.[172]

As crucial as the MMPI-2 is in assessing personality, it is not the only test used to evaluate a defendant's mental state at the time of a crime. For example, the Rogers Criminal Responsibility Assessment Scales (R-CRAS) is also frequently used.[173] Rogers et al. wanted to convert the ALI definition of insanity into 25 measurable variables, "grouped into five topics of psycho-legal relevance: organicity, psychopathology, cognitive control, behavioural control, and the reliability of the report."[174] While the test's designers reported high inter-judge reliability[175] (a high level of agreement among those who rate the test), others have been critical, noting high variability in test

---

[169] LARRY E. BEUTLER & GARY GROTH-MARNAT, INTEGRATIVE ASSESSMENT OF ADULT PERSONALITY 164-67 (2d ed., 2003). The 10 scales are: (1) hypochondriasis; (2) depression; (3) hysteria; (4) psychopathic deviate; (5) masculinity-femininity; (6) paranoia; (7) psychasthenia; (8) schizophrenia; (9) hypomania; (10) social introversion.

[170] Neal et al., *supra* note 2.

[171] *Id*.

[172] *Id*.

[173] *See* RICHARD ROGERS, ROGERS CRIMINAL RESPONSIBILITY ASSESSMENT SCALES (R-CRAS) & TEST MATERIALS (1984); RICHARD ROGERS & DANIEL W SHUMAN, CONDUCTING INSANITY EVALUATIONS (2000); Richard Rogers et al., *Assessment of Criminal Responsibility: Initial Validation of the R-CRAS With the M'Naghten and GBMI standards*, 9 INT'L J. L. PSYCHIATRY 67 (1986); Richard Rogers et al., *Evaluating Insanity: A study of Construct Validity*, 8 L.& HUM. BEHAV. 293 (1984).

[174] WRIGHTSMAN & PORTER, *supra* note 90, at 196.

[175] Richard Rogers et al., *Scientific Inquiry in Forensic Psychiatry*, 5 INT'L J. L. & PSYCHIATRY 187 (1982).

results depending on who administered it.[176] In addition to the R-CRAS, the Mental Screening Evaluation (MSE) includes questions about the general psychological history, the alleged offense, and the defendant's current mental state when the defendant takes the test.[177] However, like the R-CRAS, people criticize the MSE for its lack of standardization and formal scoring validity.[178]

This criticism outlines the first major issue with the kind of testing experts use to assess criminal defendants. On the one hand, the law has rightly determined that if a defendant possessed a mental defect at the time of the crime that prevented her from appreciating her actions, she ought not to be found criminally liable. On the other hand, assessing the defendant's state of mind rests entirely on expert witnesses who not only infer the defendant's cognitive state from surrounding circumstances and facts but also administer a battery of psychometric testing with no clear, consistent standard among the experts.

Neal and Grisso recently found that most psychologists and psychiatrists use more than one test when working within the legal system.[179] They also note a wide variation in the tests chosen, even when attempting to measure the same mental phenomenon.[180] Their findings corroborate an American Psychological Association study.[181] Perhaps even more surprisingly, King, Wade, and Tilson summarized the findings of at least 22 surveys given to forensic mental health experts.[182] The surveys asked the practitioners which tools they used in assessing patients in a legal context. The experts identified 364 distinct psychometric assessment protocols actively used or

---

[176] Stephen L. Golding & Ronald Roesch, *The Assessment of Criminal Responsibility: A Historical Approach to a Current Controversy, in* HANDBOOK FORENSIC PSYCH. 395 (I.B Weiner & A.K. Hess eds., 1987).
[177] Christopher Slobogin et al., *The Feasibility of a Brief Evaluation of Mental State at the Time of the Offense*, 8 L. & HUM. BEHAV. 305 (1984).
[178] *See generally* THOMAS GRISSO ET AL., EVALUATING COMPETENCIES: FORENSIC ASSESSMENTS AND INSTRUMENTS (2005).
[179] Tess M.S. Neal & Thomas Grisso, *Assessment Practices and Expert Judgment Methods in Forensic Psychology and Psychiatry: An International Snapshot*, 41 CRIM. JUST. & BEHAV. 1406, 1421 (2014).
[180] *Id*.
[181] *See* Wayne J. Camara et al., *Psychological Test Usage: Implications in Professional Psychology*, 31 AM. PSYCH. ASS'N 141 (2000).
[182] Neal et al., *supra* note 2, at 140.

considered appropriate for use in a court proceeding.[183] Tess M.S. Neal, Christopher Slobogin, and Michael J. Saks found that a third of the psychometric tests were not peer-reviewed. Of the ones that were, only 40% were approved for reliability and validity, while 25% were considered unreliable.[184] Perhaps most surprising was that evidentiary challenges were made in only 5% of cases in a sample conducted by the above researchers.[185] Thus, psychometric tests that are unreliable, invalid, and generally poorly represent the concepts they purport to report are objectively admitted with astounding frequency in courtrooms across the country.

Because of the inconsistencies in applying psychometric testing to assess a defendant's cognitive state at the time of the crime, as well as the inferential nature of the evidence, I argue that a scientific approach based on emerging neuroimaging technologies and mathematical modeling may help provide an objective measure of the defendant's state of mind at the time of a crime.

VIII. IMAGING DYNAMIC BRAIN NETWORKS: AN INTRODUCTION

Before looking at how neurotechnologies are used to assess a defendant's *mens rea*, it is instructive to start with an overview of the available technologies and a description for the non-specialist of how they work.

In 1879, Angelo Mosso became the chair of Physiology at the University of Turin. While there, Mosso did some exciting work measuring brain pulsations in patients with skull defects.[186] Mosso based his experiments on his prior work exploring a unique phenomenon in newborn infants.[187] The blood inside the skull increases as the heart pumps blood to the brain, causing the supple newborn skull to distend slightly.[188]

---

183 *Id.*

184 *Id.* at 164.

185 *Id.*

186 Lewis Hong, *Angelo Mosso*, LIFE IN THE FASTLANE (Nov. 10, 2020), https://litfl.com/angelo-mosso/#:~:text=1879%20%E2%80%93%201904%20%E2%80%93%20Chair%20of%20Physiology%20at,Awarded%20National%20Fellow%20of%20the%20Accademia%20dei%20Lincei.

187 *Id.*

188 *Id.*

Blood drains from the brain when the heart relaxes, causing the skull to retract.[189] Mosso converted these rhythmic distensions and retractions of the newborn skull into traceable waves that allowed him to calculate how much the blood volume changed in each cardiac cycle.[190] Because Mosso's adult patients displayed defects in their skulls, he could measure the distension/contraction in adults as quickly as he could in a newborn.[191]

On their face, Mosso's experiments show high scientific relevance, although their real-world applicability is not readily apparent. In a stroke of brilliant insight, Mosso refined his experiment and had his subjects perform different tasks while he measured their skulls.[192] As a result, he noted something that would forever change the future of physiology and brain science.[193] If he asked a patient to do a more complex mental task, like solve a math problem, their brain's pulsations increased.[194] This pulsation spike meant that increasing mental activity leads to increased blood to the brain.[195] This notion seems intuitive to us and is based on how we experience this physiological maxim throughout our bodies: more blood flows to areas that need it. Furthermore, showing a similar increase in blood flow when the brain works harder makes Mosso's initial findings so groundbreaking.

Mosso further refined his theories and experimental apparatus after his appointment at Turin. One of the biggest problems was that Mosso could not test his theory in healthy subjects because brain pulses are only readable if the skull is either immature or defective.[196] To solve the problem, he designed and built a table that would be carefully balanced but could tip either towards the head or the feet.[197] Mosso would ask his subjects to lie on the table, balance them, and then ask

---

[189] *See generally* Stefano Sandrone et al., *Weighing Brain Activity with the Balance: Angelo Mosso's Original Manuscripts Come to Light,* 137 BRAIN 621 (2013).
[190] *Id.*
[191] *Id.*
[192] *Id.*
[193] *Id.*
[194] Marcus E. Raichle, *A Brief History of Human Brain Mapping*, 32 TRENDS NEUROSCIENCES 118 (2009).
[195] *Id.*
[196] Sandrone, *supra* note 189.
[197] *Id.*

them to perform cognitive tasks.[198] If the task required more thinking, more blood would flow to the brain, increasing the weight of the table at the head side, thus tipping the balance.[199] As blood left the brain when there was no cognitively demanding task, it returned to the core and legs, and the table regained balance.[200] Excitingly, in the late 19th century at Turin, Mosso's table became the first functional[201] neuro-measuring device.[202]

In the late 1930s, a physicist named Isidor Rabi first showed that water, any water, has a unique property: the nuclei of the atoms that make up tiny molecules that make up a glass of water can get excited when you put them near a magnetic field.[203] The discovery was exciting for physicists, but it was not until the early 1970s that it showed its actual value. Paul Lauterbur published the first image of the excited water molecules in *Nature*.[204]

Let us picture it this way: water molecules are teenagers on the dancefloor at a high school dance under normal conditions. Before the music starts, they may walk around, they may stand to the side, they may run to meet with their friend, and others cannot wait to leave. The movement of the teenagers is random at this point. Now, imagine the DJ putting on the latest hit, that one song that every teenager loves and every parent dreads. Upon hearing their favorite song, the students all begin to dance to it; they all rush to the dance floor and cannot stop dancing to the rhythm of the music. When the music goes quiet, they return to their random moving about like before. In this analogy, the teenagers are the water atoms, and before the music plays, they are all floating around happily in their little glass. The music is the magnet. As soon as they hear

---

198 *Id*.

199 *Id*.

200 *Id*.

201 As contrasted with "structural" imaging. Functional imaging allows us to see the organ do its job, while structural imaging lets us see what the organ looks like.

202 *See* Stefano Sandrone et al., *Angelo Mosso (1846–1910)*, 259 J. NEUROLOGY 2513 (2012).

203 Isidor I. Rabi et al., *A New Method of Measuring Nuclear Magnetic Moment*, 53 PHYSICAL REV. 318 (1938).

204 Paul C. Lauterbur, *Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance*, 242 NATURE 190 (1973).

that song, especially one *they* like, they all start to dance in sync and to the rhythm. When the music turns off, they return to their old behavior. Now, Isidor Rabi first discovered that the water atoms dance to the music of the magnetic field. Paul Lauterbur showed that when atoms get excited, one can take a picture of them; in other words, the act of excitation produces a measurable signal that disappears when the atoms are no longer excited.[205]

Lauterbur performed these initial experiments using plain water in test tubes.[206] However, his findings are interesting because water is also the most abundant substance in our bodies, especially in our brains. Specifically, because we have so much water in our bodies, all we must do is put the part of the body we are interested in studying near a large enough magnet, and with the suitable receiver, we can imagine the water dancing to the rhythm of the magnet. This notion portrays the idea behind using Magnetic Resonance Imaging (MRI) to see inside the skull and look at the brain.

Angelo Mosso was the first to measure blood flow and volume changes in the brain in response to increased cognitive work.[207] MRI, in turn, allows a special magnet and its receiver to excite and then measure the position of water atoms as they float inside the brain so that a physician or researcher can take a picture of the brain.[208] In the early 1990s, several researchers tried to see if they could use the blood oxygen level-dependent (BOLD) response to generate a picture of what is happening inside the skull.[209] The BOLD response measures how much oxygen is present in the blood at any given time in an area of the brain.[210] One of the blood's primary purposes is to deliver

---

[205] *Id.*

[206] *Paul Lauterbur*, MAGNET ACADEMY, https://nationalmaglab.org/education/magnet-academy/history-of-electricity-magnetism/pioneers/paul-lauterbur (last visited Sept. 11, 2022).

[207] Sandrone, *supra* note 189.

[208] *MRI*, MAYO CLINIC, https://www.mayoclinic.org/tests-procedures/mri/about/pac-20384768 (last visited Oct. 8, 2022).

[209] *See* Richard B. Buxton, *The Physics of Functional Magnetic Resonance Imaging*, 76 REPS. ON PROGRESS PHYSICS 1 (2013).

[210] Kenneth K. Kwong et al., *Dynamic Magnetic Resonance Imaging of Human Brain Activity During Primary Sensory Stimulation*, 89 PROC. NAT'L ACAD. SCIS. 5675 (1992).

oxygen to body parts that need it.[211] We have already discussed how Mosso's studies showed that more blood flows to the brain while it does more work. This finding makes sense because the brain cells need more oxygen while working harder. Nevertheless, overall blood flow to the brain – as a measure of volume change – does not tell us which brain parts are working harder. Just like a runner has more blood going to their legs when running than to their arms, parts of the brain doing the cognitive work will need more blood than other parts with other responsibilities. This is where the BOLD signal entered the picture, and the scientific world never looked back. Now, instead of measuring the overall blood, or the water inside the brain, we can take a picture in real-time of where the blood is *going* inside the patient's brain. For example, a person shown a beautiful painting while in the scanner will show more blood flowing to the part of his brain that processes visual information than to the part that processes sound.

Conversely, a person who listens to a piece of music while blindfolded will show a dramatic increase in blood to the part of the brain that processes sound and almost no change in blood to the visual part. The beauty of this kind of information is that it can be infinitely varied to see how the brain reacts. What if we put an ugly painting instead of a beautiful one? Or, what if the painting lacked color? Or what if we played one piece of music in one ear and another in the other ear? These variations allow scientists to study the brain's activity in response to stimuli. Indeed, this is the basis of functional Magnetic Resonance Imaging (fMRI). It is termed functional because it takes a real-time picture of the brain at work – how it functions – and Magnetic Resonance Imaging because it measures the excitation of atoms in response to a magnetic field.[212]

If I want to see which part of the brain becomes active when I ask a person to tap his fingers, I can put the person into an MRI scanner, start the scan, and ask him to do that task. As he gently taps his fingers, neurons, or brain cells, in the part of his brain that controls movement will start working harder than usual. As he continues to tap his fingers, those neurons need more resources to keep working because they have depleted

---

[212] *See* Nikos K. Logothetis et al., *Neurophysiological Investigation of the Basis of the fMRI Signal*, 412 NATURE 150 (2001).

their stores. So, the cells cleverly send a signal to the nearby blood vessels for more oxygen and glucose, which are the building blocks of their energy. The nearby blood vessels receive the signal and immediately deliver a fresh batch of blood so the neurons can take what they need to keep working. We can imagine this delivery of fresh blood, which we call the BOLD response. Furthermore, we can now see the location of those cells because once we superimpose the activity map onto a brain map, we see precisely where the demanding brain cells are in response to the tapping of the finger. This subtype of functional neuroimaging is task-based because we begin with a task and then follow the blood.[213]

However, what if one wanted to measure brain activity when the patient is not engaged in a task? Or if the patient is thinking or forming goals before initiating a task? Resting-state fMRI (rs-fMRI) is a technique used in brain mapping, showing how two parts of the brain talk to each other when the patient is not performing an explicit task.[214] Rs-fMRI is a powerful tool because it can measure intrinsic fluctuations in brain activity by looking at how much blood is flowing to different parts and then correlate those fluctuations with each other, thus forming a functional map of the brain.[215]

For example, imagine if a space alien came to earth without knowing how phones work. Our friend, the space alien, is fascinated with human behavior and observes two people. The two people are separated by a considerable distance, making the same motion with their arms: they reach out, pick up a device, put it up to their heads, gesticulate with their faces for a while, then both put the device down at the same time. The space alien finds this behavior curious and eventually realizes that these actions are not random events but are related. Even more interestingly, the simultaneous raising of the hands occurs because the two people *communicate*, though separated by distance. In this example, we are the space alien when observing the brain, a landscape foreign to us in which we do not speak the language or understand the local

---

[213] *Id.*

[214] *See* Bharat B. Biswal, *Resting State fMRI: A Personal History*, 62 NEUROIMAGE 938 (2012); Randy L. Buckner et al., *Opportunities and Limitations of Intrinsic Functional Connectivity MRI,* 16 NATURE NEUROSCIENCE 832 (2013).

[215] *Id.*

habits. Resting-state fMRI allows us to see behaviors – in this case, small fluctuations in blood flow – between regions separated by a distance, correlate those behaviors, and then realize who is communicating with whom inside the brain.[216]

IX. GRAPH THEORY: AN INTRODUCTION TO THE MATHEMATICS OF NETWORKS AND THEIR APPLICABILITY TO THE BRAIN

This idea of connectivity between brain regions is relatively new in neuroscience, although its theoretical framework arose in a fascinating mathematics thought experiment. On August 25, 1735, Leonard Euler presented a paper containing a solution to one of the most troubling mathematical problems that remained unsolvable until that time. Allow me to give a brief background to the problem.

The quiet town of Königsberg in Prussia sits on the banks of the Pregel River. The river divides the city into four regions connected by seven bridges. Local legend had it that the city's citizens would walk around every Sunday afternoon to take in the beautiful sights. Along their walks, they would devise a clever game to play, in which the object was to walk around the entire city, crossing each of the seven bridges only once.[217] This task proved far more complex than anyone imagined, yet no one could understand why that might be. These were the events leading up to Euler's groundbreaking discovery. I shall discuss shortly the significance of this bridge problem to the functioning of the brain and the emergence of cognition, intent, purpose, and even the insanity defense itself.

Euler was a highly prolific mathematician, writing a paper a week, so he found the bridge problem trivial. To solve the problem, Euler referred to a mathematical concept of *geometria situs* first coined by Gottfried Leibniz.[218] *Geometria situs* is the geometry of position.[219] The basic idea behind this geometry of position is topology, that is, those properties of any geometric setup that remain constant under deformations like

---

[216] *Id.*

[217] *See generally* NORMAN BIGGS ET AL., GRAPH THEORY (1986).

[218] Evgeny Zaytsev, *Euler's Problem of Königsberg Bridges and Leibniz' Geometria Situs*, 58 ARCHIVES INTERNATIONALES D'HISTOIRE DES SCIENCES 151 (2008).

[219] *Id.*

stretching, compressing, twisting, or crumpling.[220] In solving the Königsberg bridge problem, Euler argued it did not matter how long the bridges were or how far apart they were from each other; instead, what matters is how they *connected* to the stable points – the islands. Upon first glance, this discovery may appear meaningless for anyone outside of the field of mathematics, yet its implications are far-reaching. Graph theory was born from Euler's proof, a unique mathematical field that studies network topology – that is, a network's physical and logical layout.

A graph is a network; every network contains stable islands or nodes, and connections between the nodes – the bridges – called edges.[221] Thus, graphs are visual representations of networks and can plot and predict the changes across the network as the connections between the constituent parts change. Let us pretend for a moment that a graph is like the aviation industry. Hubs, or airports, are scattered throughout a geographic area like the United States. Every hub processes thousands of incoming units – in this example, airplanes – every day and remains stationary within the network. Furthermore, every hub is connected to the others by the airplanes flying between them. This connectivity between airports is a dynamic system because it constantly changes. For example, on some days of the week, LaGuardia Airport might have higher incoming traffic from Los Angeles International Airport (LAX) than does, say, Chicago O'Hare International Airport.

On the other hand, Detroit Metro Airport might have the lowest incoming traffic from Boston Logan International Airport on Thursday evenings. The system adapts and changes in real-time, yet it does not do so in a chaotic way. Instead, it is incredibly well-organized. The exciting thing about a graph is that features that seem to occur randomly, in fact, have underlying causes. For example, in our airport analogy, at first glance, it might appear to be a fluke that commercial flights leaving from LAX for John F. Kennedy Airport (JFK) on Fridays are usually delayed by 20 minutes. However, graph theory allows us to see how those connections between LAX and JFK

---

[220] *Id.*

[221] *Id*. *See also* CLAUDE BERGE, THÉORIE DES GRAPHES ET SES APPLICATIONS (1958).

flights might be influenced by an unknown variable, like a flight delay from Dallas Fort Worth International Airport. Thus, changes or perturbations to the network have downstream effects on the connectivity between other parts of the network. At the same time, it adjusts to those changes to regain equilibrium, which can mean either returning to its previous state or a new one.

How does this digression into mathematics and airplanes apply to the thesis of this paper? The answer lies in neuronal topology – that is, how the brain is organized into static and connective regions that exhibit functional connectivity – dynamic changes in their connectivity depending on the brain's state. Let us analogize to the airport: the hubs, or airports, are discrete, definable parts of the brain that do a specific task. They may receive incoming information and process it or send commands to other parts of the brain or body – like airplanes landing or taking off from one airport and going to another. Furthermore, like airports, these brain regions do not exist in a vacuum; instead, by being connected to other areas through a complex web, they can modulate their responses to what the network is doing. Just like perturbation in one end of the country – say, a thunderstorm in Miami, may influence the volume of air traffic to Minneapolis.[222]

Above, I discussed how goal-directed behavior is generated in one area of the cortex, which then communicates that goal to the primary motor cortex, generating voluntary action to accomplish the goal. In this way, we can see a simple network emerge where there are nodes or areas of the brain responsible for particular aspects of the voluntary action sequence. However, their activity is modulated because of their interconnectedness with other areas in the pathway. Thus, goal-directed behavior, or what the law would call "conscious-object," likely arises due to the functional interplay between multiple distinct regions in the prefrontal cortex discussed below. These areas draw on sensory, emotional, and memory information stored in other parts of the brain and form the goal or motive the action attempts to satisfy.

As I have shown, goals and motives arise in the prefrontal cortex, the area of the brain located directly behind

---

[222] *See generally* Caroline Alionte et al., *From Symmetry to Chaos and Back: Understanding and Imaging the Mechanisms of Neural Repair after Stroke*, 288 LIFE SCIS. 120161 (2022).

the forehead.[223] Studies show that defined areas within the prefrontal cortex generate hypothetical outcomes to known situations. In particular, the dorsolateral prefrontal cortex (DLPFC) is involved in putting together our various sensory perceptions of the world and producing such hypothetical outcomes.[224] Furthermore, cells within the DLPFC modulate their activity depending on the "outcomes expected or obtained from specific action."[225] In addition to the DLPFC, Kenji Matsumoto, Wataru Suzuki, and Keiji Tanaka looked at other areas in the frontal lobe that are activated when choosing to act. In their experiments, they trained monkeys to complete a task in which the monkeys could choose an action from a set of reasonable actions based on a goal.[226] While the monkeys selected the action, the researchers recorded the activity of cells in the prefrontal cortex.[227] They found increased activity in the medial and lateral parts but not in other areas.[228]

---

[223] Maurizio Corbetta & Gordon L. Shulman, *Control of Goal-Directed and Stimulus-Driven Attention in the Brain*, 3 NATURE REV. NEUROSCIENCE 201 (2002).

[224] *See* Vivek Prabhakaran et al., *Integration of Diverse Information in Working Memory Within the Frontal Lobe*, 3 NATURE NEUROSCIENCE 85 (2000); Alan Baddeley, *The Episodic Buffer: A New Component of Working Memory?*, 4 TRENDS COGNITIVE SCI. 417 (2000); Karen J. Mitchell & Marcia K. Johnson, *Source Monitoring 15 Years Later: What Have We Learned From fMRI About the Neural Mechanisms of Source Memory?*, 135 PSYCH. BULL. 638 (2009).

[225] Se*e generally* Hiroshi Abe & Daeyeol Lee, *Distributed Coding of Actual and Hypothetical Outcomes in the Orbital and Dorsolateral Prefrontal Cortex*, 70 NEURON 731 (2011); Masataka Watanabe, *Reward Expectancy in Primate Prefrontal Neurons*, 382 NATURE 629 (1996); Matthew I. Leon & Michael N. Shadlen, *Effect of Expected Reward Magnitude on the Response of Neurons in the Dorsolateral Prefrontal Cortex of the Macaque*, 24 NEURON 415 (1999); Kenji Matsumoto et al., *Neuronal Correlates of Goal-Based Motor Selection in the Prefrontal Cortex*, 301 SCI. 229 (2003); Dominic J Barraclough et al., *Prefrontal Cortex and Decision Making in a Mixed-Strategy Game*, 7 NATURE NEUROSCIENCE 404 (2004); Hyojung Seo & Daeyeol Lee, *Behavioral and Neural Changes after Gains and Losses of Conditioned Reinforcers*, 29 J. NEUROSCIENCE 3627 (2009).

[226] Kenji Matsumoto et al., *Neuronal Correlates of Goal-Based Motor Selection in the Prefrontal Cortex*, 301 SCI. 229 (2003).

[227] *Id.*

[228] *Id.*

Moreover, the orbitofrontal cortex (OFC) also plays a role in adjusting behavior to meet a goal. It is guided in this behavioral adjustment by the hypothetical goal outcomes that the DLPFC and medial prefrontal cortex generate.[229] We can think of these anatomically defined areas of the prefrontal cortex as airports that communicate to each other, and out of their communication emerge the pattern of higher cognitive states like purposeful action. Therefore, because of this communication or functional connectivity among the centers, intentionality likely arises, not in any one area acting alone. The medial, lateral, dorsolateral, and orbitofrontal cortices all work together, modulating each other's activity and ultimately forming a goal out of this connectivity. The goal's formation relies on healthy connections with other regions that ensure it is based on sound reasoning and societally acceptable. In their totality, these brain regions form a network topology, much like the bridge problem Euler solved several centuries ago. First, the goal gets sent to the areas of the motor cortex from which the best movement plan is selected, thus producing the voluntary action which will fulfill the goal. From there, the motor cortices send motor impulses to the spinal cord and to the groups of muscles that then execute the action.

A graph theory approach to neuroscience is appealing because it considers the interactions between the network nodes, not the nodes acting alone. Just as the national aviation system emerges as an organic, dynamic system of the movement of information between hubs, so too does the prefrontal cortex shuttle information between the centers listed above; it is from this dynamic movement where higher cognition, like a motive, goal, or intent, likely emerges.[230]

The airport analogy helps us visualize this idea, but another helpful metaphor is the global economy. Economic trends emerge out of the interaction of billions of people engaging in commerce at the micro-level. Each person within the network buys, sells, or generally contributes to the

---

[229] Nathalie Camille et al., *The Involvement of the Orbitofrontal Cortex in the Experience of Regret*, 304 SCI. 1167 (2004); Giorgio Coricelli et al., *Regret and Its Avoidance: A Neuroimaging Study of Choice Behavior*, 8 NATURE NEUROSCIENCE 1255 (2005).

[230] *See* Ian M. McDonough & Kaoru Nashiro, *Network Complexity as a Measure of Information Processing Across Resting-State Networks: Evidence from the Human Connectome Project*, 8 FRONTIERS HUM. NEUROSCIENCE 1 (2014).

economy. When viewed as a whole, the tiny economic interactions get amplified globally such that a pattern emerges independent of the action of any person who is a part of it. The network is resilient to minor disruptions, like if one local business closes. However, as disruptions become more prevalent, more businesses close, and more people lose their jobs. Then, the effects on the global economy become more noticeable until there is a tipping point that the network sets on its own, and once this threshold is crossed, fundamental changes in the emergent trends.[231] Graph theory allows us to study these emergent, dynamic properties of any network, make predictions about the network's future states, and try to find the source of any disruptions within the network.

X. DISRUPTED BRAIN NETWORKS: GRAPH THEORY OF PSYCHOLOGICAL DISORDERS AND CRIMINAL INSANITY

The legal concept of purposeful action is likely an emergent property arising out of the interaction of billions of brain cells that make up critical regions in the prefrontal cortex, like the medial, dorsolateral, and orbitofrontal cortex.[232] Yet, what happens when we observe an abnormal or unexpected emergent property? We know from experience that once a person intends to complete an action, no matter how rudimentary the action might be, the process works seamlessly from the intent's conception to the act's execution. Reaching across the table for a glass of water and raising it to my mouth occurs within milliseconds of me deciding I want water. However, any breakdown in either the functional connectivity from which the goal for water gives rise or the goal's communication downstream for the correct selection of the voluntary motor plan becomes quite evident to an observer. Suppose my goal is to get a sip of water, and I attempt to stretch out my right arm, and instead, my left arm goes out. In that case, there is something wrong with the internal system. Alternatively, if the arm extends but the fingers cannot grasp the glass, no matter how hard I try, there is an apparent breakdown among the motor pathways.

---

[231] *See* Doaa T. Taha et al., *Measurement for a Nonlinear Dynamical Theory of Acute Cell Injury*, 110 BIOPHYSICAL J. 479a (2016).

[232] *See generally* STEVEN JOHNSON, EMERGENCE: THE CONNECTED LIVES OF ANTS, BRAINS, CITIES, AND SOFTWARE (2002).

Nevertheless, the breakdown may be more subtle. Perhaps, as outsiders, we perceive the action as voluntary, yet there was a subtle breakdown in the prefrontal cortical network giving rise to the goal. This disconnect is, of course, the crux of the criminal insanity defense. Behavior appears to be the ordinary emergence from the internal network. Nevertheless, minor abnormalities in the network prevent proper goal formation or understanding the consequences of the goal. Alternatively, minor network alterations lead to an abnormal goal.

Disordered brain connectivity is becoming a more prevalent theory of mental illness, especially as it applies to goal and intent formation and action execution. For example, Lefort-Besnard et al. argue that in schizophrenia, "abnormal connectivity between large-scale networks and the DMN can provide insight into the longstanding 'disconnection hypothesis' that explains schizophrenia pathophysiology as coupling impairments due to context-dependent synaptic modulation."[233] They continue and say:

> Such disconnection of large-scale networks may contribute to positive symptoms through the failure of attentional reallocation and monitoring processes, but also to cognitive symptoms through impaired perceptual inference and disturbance of associative learning, as well as to negative symptoms due to inability of learning from and adapting to social environments. Together, these converging lines of evidence highlight that coupling patterns of canonical networks and the DMN may service as an important biomarker for many aspects of the psychopathology of schizophrenia.[234]

This contention suggests that many of the cognitive abnormalities observed in a mental disorder have underlying

---

[233] *See* Jérémy Lefort-Besnard et al., *Different Shades of Default Mode Disturbance in Schizophrenia: Subnodal Covariance Estimation in Structure and Function*, 39 HUM. BRAIN MAPPING 644 (2017); Karl J. Friston*, The Disconnection Hypothesis,* 30 SCHIZOPHRENIA RSCH. 115 (1998).

[234] *Id.*

biological causes that are rooted in the way different parts of the brain share information with each other. It is like a permanent thunderstorm in Miami that delays all flights forever out of LAX, which then indirectly affects the timing and efficiency of all domestic U.S. flights. A domino effect of interconnective dysfunction may result in psychopathology. Stéphanie Lefebvre et al. note:

> It appears that an isolated brain lesion or dysfunction cannot be responsible for the whole clinical spectrum of mental disorders (in contrast to models for stroke or other common neurological diseases) and especially cannot explain the intermittence of the occurrence of hallucinations, which might be better explained by aberrant integration and connectivity between large-scale networks.[235]

Furthermore, rs-fMRI allows the researcher or clinician to see minute changes in blood flow to the nodes that make up networks of the brain that, through working together, give rise to emergent cognitive features like purposeful action. Lefebvre et al. contend that "the study of functional connectivity with resting-state functional magnetic resonance imaging (rs-fMRI) has revealed that impairment in the interaction between several intrinsic connectivity networks (ICNs) could be related to specific psychopathological processes."[236]

---

[235] Stéphanie Lefebvre et al., *Network Dynamics During the Different Stages of Hallucinations in Schizophrenia*, 37 HUM. BRAIN MAPPING 2571 (2016).

[236] *Id*; *see also* Karl J. Friston*, The Disconnection Hypothesis,* 30 SCHIZOPHRENIA RSCH. 115 (1998) (Network dysfunction does not only apply in cases of schizophrenia); Scott Tillem et al., *Psychopathy Is Associated With Shifts in the Organization of Neural Networks in a Large Incarcerated Male Sample*, 24 NEUROIMAGE CLINICAL 1, 10 (2019) (finding "[a]t the level of overall brain organization, it appears that subcortical structure (e.g., amygdala, caudate, and hippocampus), collectively, act as less of a central hub in the global flow of information throughout the brains of individuals higher in psychopathy. More specifically, the current findings indicate that, when engaging in spontaneous, undirected cognition (i.e., at rest), less information flows through subcortical structures in

Thus, rs-fMRI paints a picture of the brain's default network in physiological and pathophysiological conditions precisely because it can measure temporally correlated fluctuations in blood flow to various brain areas in real-time. Disruptions in brain networks manifest themselves as changes or even breakdowns in functional connectivity between the nodes. That is, the blood flow may show a more chaotic pattern of coactivity when compared to a healthy control brain. This chaotic pattern, in turn, may disrupt the brain's ability to form appropriate goals. When a goal does emerge, it elicits a behavioral response, however voluntary, that is not in congruence with that of a reasonable, healthy person under the same circumstances.[237] Indeed, diseased cognition may generate an intent that, while present, is based on faulty reasoning. This intent, in turn, generates voluntary action that aims to satisfy the delusional cognition that motivates it.

In other words, the rs-fMRI of brain networks allows us to generate a "brain print," which, much like a fingerprint, identifies the individual and also determines any inherent flaws in the flow of information that present themselves as impaired executive functioning and appreciation of the consequences of one's actions.[238] In this way, graph theory and functional neuroimaging allow us to keep the legal categories of criminal insanity and *mens rea* distinct but not separate.

XI. CONCLUSION

I began this piece with a hypothetical situation in which a criminal defendant claims, at the present time, to be suffering a psychotic episode that began in the past, which prevents him from forming goals that are congruent with the situation. After killing his wife, the defendant argues he should be found not guilty because of criminal insanity. Accordingly, psychiatrists evaluate him and administer the standard battery of

---

psychopathic individuals. This suggests that spontaneous neural communication between cortical and subcortical structures, as a whole, may be disrupted in psychopathy.").

[237] *See* Alessandra Griffa et al., *Brain Connectivity Alterations in Early Psychosis: From Clinical to Neuroimaging Staging*, 9 TRANSLATIONAL PSYCHIATRY 1 (2019) (correlating functional neuroimaging with clinical staging in the development of psychotic episodes).

[238] *See generally* Victor Vicens et al., *Structural and Functional Brain Changes in Delusional Disorder*, 208 BRIT. J. PSYCHIATRY 153 (2016).

psychometric tests to assess his state of mind. Yet, both his defense counsel and the prosecutor remain skeptical about whether he is being truthful. Naturally, the defense will argue for insanity regardless, and both sides will present expert witnesses who will provide data and the results of their evaluations. Nevertheless, the fundamental problem is that the experts cannot always agree on the best tests to use, each relying on their combination of tests, thus reducing the measure's reliability. Consequently, both sides will likely arrive at different conclusions about the defendant's mental state at the time of the murder, thus putting the burden on the fact-finder to decide.

The MPC attaches the highest culpability to a defendant who acted with purpose or a conscious object. A conscious object is synonymous with a goal or motive which arises out of the neuronal topology or connective interplay of anatomically defined regions in the prefrontal cortex. Furthermore, once it emerges, the brain communicates it to the parts that select the voluntary action motor plan to achieve the goal. Disruptions in the prefrontal cortex likely lead to improper goal selection, which presents as psychopathologies. Once formed, an improper goal communicates a voluntary action, and the body executes it. Thus, externally the system presents with internal cohesion, yet there is a pathological breakdown in the goal-production network. Fortunately, graph theory, with the help of rs-fMRI, can help identify these changes in the work of a diseased brain and make the task of the fact finder in a criminal case easier by providing an objective, physiological brain print of the defendant's ability to form goals and appreciate their consequences.