

MINDFULNESS, EMOTIONS, AND ETHICS: THE RIGHT STUFF?

Ellen Waldman*

I. INTRODUCTION

What role do emotions play in ethical decision-making? Philosophers have long debated the question, disagreeing about both the nature of “the good” and how best to achieve it. Rationalists ground one’s capacity for virtue in logic and deliberate cognition, while moral intuitionists look to one’s capacity for feeling deeply. Immanuel Kant, for example, maintained that right conduct flowed from a sense of duty that functioned independently of emotion.¹ Conversely, David Hume argued that all right action involved sentiment and that reason, stripped of passion, could not impel ethical choice.²

Philosophers are not alone in their fascination with the question. Psychologists also have delved into the relationship between emotion and moral development, creating varying models of maturation that either embrace or reject emotion as a critical component of moral discernment. The competing visions of theorists Leonard Kohlberg and Carol Gilligan illustrate the divide: Kohlberg’s schema of moral development viewed linear, abstract thinking as the apogee of moral sophistication, whereas Gilligan gave feelings of relatedness, empathy, and affiliation pride of place.³

Today, debates in the “soft sciences” of the mind spill into the “hard sciences” of the body.⁴ Interest in the biological bases of emotion invigorates neuroscience, and developments in functional magnetic resonance imagery (fMRI) promise methods for mapping the synaptic pathways that induce affective states.⁵ Although we can now detect activity in portions of the brain associated with emotional experience, it remains unclear whether those electrical surges push us in “right” or “wrong” directions.

* LL.M., University of Virginia in Mental Health Law, with an emphasis in Conflict Resolution; J.D., New York University. Thanks go to my research assistant Laura Manteghian, alumna of Quinnipiac Law School, for her prodigious efforts tracking down stray texts and creating order of a footnoting kind – as well as her gracious good-humor and spirit.

¹ IMMANUEL KANT, *CRITIQUE OF PRACTICAL REASON* 110 (H. W. Cassirer, G. Heath King & Ronald Weitzman eds., Cambridge Univ. Press 2007) (1788).

² 1 DAVID HUME, *A TREATISE OF HUMAN NATURE* 367-93 (David Fate Norton & Mary J. Norton eds., Oxford Univ. Press 2007) (1739).

³ See CAROL GILLIGAN, *IN A DIFFERENT VOICE* (1982); Lawrence Kohlberg, *The Claim to Moral Adequacy of a Highest Style of Moral Judgment*, 70 *J. PHIL.* 630, 630-36 (1973).

⁴ See, e.g., Daniel Langleben et al., *Telling Truth from Lie in Individual Subjects with Fast Event-Related fMRI*, 26 *HUM. BRAIN MAPPING* 262 (2005) (examining correlations between increased prefrontal and parietal brain activity and deception).

⁵ K. Luan Phan et al., *Functional Neuroanatomy of Emotion: A Meta-Analysis of Emotion Activation Studies in PET and fMRI*, 16 *NEUROIMAGE* 331, 331-48 (2002).

In the mediation world, scholars and practitioners frequently treat emotion as the unruly step-child of the problem-solving mind.⁶ They warn that untamed emotion can provoke outbursts, distort interpretation, impair effective listening, and scuttle Pareto-optimizing⁷ moves in countless other ways.⁸ For these reasons, leading mediation theorists present emotion as a force to be blunted,⁹ manipulated,¹⁰ or leveraged¹¹ in the service of “getting to yes.”¹²

In his most recent work, Professor Leonard Riskin similarly characterizes emotion as a potential negotiation saboteur and offers “mindful practice” as a useful corrective.¹³ Specifically, Professor Riskin argues that mindful mediation can help negotiators gain better control over their wandering minds and negative emotions, and achieve more satisfying, interest-based solutions.¹⁴ In an earlier work, Riskin touted mindful practice as a way to encourage emotional awareness, discernment, and detachment, as well as a spirit of intercon-

⁶ See generally Don Ellinghausen Jr., *Venting or Vipassana? Mindfulness Mediation's Potential for Reducing Anger's Role in Mediation*, 8 CARDOZO J. CONFLICT RESOL. 63, 64 (2006) (discussing contemporary views of anger as a destructive, unjustified element that deprives negotiators of intelligent choice).

⁷ A Pareto optimal move or a Pareto improvement is a change “to a deal that make[s] at least one person better off without making anyone worse off.” DEEPAK MALHOTRA & MAX H. BAZERMAN, *NEGOTIATION GENIUS: HOW TO OVERCOME OBSTACLES AND ACHIEVE BRILLIANT RESULTS AT THE BARGAINING TABLE AND BEYOND* 65 (2007).

⁸ See Daniel L. Shapiro, *Preempting Disaster: Pre-Mediation Strategies to Deal with Strong Emotions*, in *THE BLACKWELL HANDBOOK OF MEDIATION: BRIDGING THEORY, RESEARCH, AND PRACTICE* 309, 310-25 (Margaret S. Herrman ed., 2006).

⁹ DOUGLAS STONE, BRUCE PATTON & SHEILA HEEN, *DIFFICULT CONVERSATIONS: HOW TO DISCUSS WHAT MATTERS MOST* 89-91 (Penguin Books 2000) (1999).

¹⁰ See ROBERT B. CIALDINI, *INFLUENCE: THE SCIENCE OF PERSUASION* (2001). Many trainers and practitioners draw on Cialdini's “principles” as part of their efforts to coax recalcitrant disputants toward settlement. For evidence of the impact of Cialdini's work on the mediation field, see James R. Coben, *Gollum, Meet Sméagol: A Schizophrenic Rumination on Mediator Values Beyond Self-Determination and Neutrality*, 5 CARDOZO J. CONFLICT RESOL. 65, 75-76 (2004).

¹¹ ROGER FISHER & DANIEL SHAPIRO, *BEYOND REASON: USING EMOTIONS AS YOU NEGOTIATE* 99 (2005).

¹² See ROGER FISHER & WILLIAM URY, *GETTING TO YES: NEGOTIATING AGREEMENT WITHOUT GIVING IN* 11 (Bruce Patton, ed., 2d ed. Houghton Mifflin Harcourt 1992) (1981) (“[H]uman beings are not computers. We are creatures of strong emotions who often have radically different perceptions and have difficulty communicating clearly. Emotions typically become entangled with the objective merits of the problem. Taking positions just makes this worse because people's egos become identified with their positions. Hence, before working on the substantive problem, the ‘people problem’ should be disentangled from it and dealt with separately.”).

¹³ See Leonard L. Riskin, *Further Beyond Reason: Emotions, the Core Concerns, and Mindfulness in Negotiation*, 10 NEV. L.J. 289 (2010).

¹⁴ Riskin defines mindful awareness as a “certain way of paying attention . . . with equanimity and without attachment—to whatever passes through the conventional sense (touch, taste, smell, vision, hearing) and the mind (thoughts).” *Id.* at 308. He maintains that practicing mindfulness induces a stance of “kindly curiosity,” which “allows the person to investigate whatever appears without falling prey to automatic judgments or reactivity.” *Id.* at 321 n.59. Maintaining this stance will, in Riskin's view, help negotiators stick to Roger Fisher and Daniel Shapiro's Core Concerns System when self-centeredness, negative emotions, and distractedness threaten to intervene.

nectedness and empathy.¹⁵ These traits, he argued, would help mediators recognize temptation and avoid unethical boundary-crossings.¹⁶ The link in both articles is the claim that mindful mediation helps individuals guard against self-centeredness and self-destructive, emotionally-impelled aggression. Additionally, each article urges mindful contemplation as a means of fostering the equanimity and expansiveness central to effective negotiation and ethical mediation.¹⁷

I do not take genuine issue with Riskin's arguments. Being able to experience provocation and the accompanying rise in blood pressure from a detached attitude of "gentle loving-kindness" seems like a good thing from any number of perspectives. Instead, I examine in more detail the relationship Riskin posits between mindful practice and ethical decision-making. Does acquiring "present-moment, non-judgmental awareness" in daily life lead to better ethical deliberation?¹⁸ Can it help us pay better attention to the "cautionary thoughts" that flash through our minds as we contemplate problematic behavior?¹⁹ Does our moral sense evolve from careful thinking, or do the emotional intuitions that mindfulness makes salient play a larger role?

This essay celebrates Riskin's call to arms while suggesting some limits to what mindfulness can achieve in the ethical realm. I discuss recent developments in neuroethics that imply a prominent role for emotions in establishing ethical restraint.²⁰ I also survey a growing body of evidence that suggests the directive power of our emotions remains largely hidden from and impervious to the control of our "reasoning" selves.²¹ Lastly, I examine what Riskin has, in an earlier work, described as the ethical hard case²² in light of recent explorations into the emotional wellsprings of deontological versus consequentialist thinking.²³ Although the mediation community need not wade deeply into the debates currently roiling social psychologists, it is useful to reflect on the genesis of our ethical commitments and whether they continue to serve the field's long-term goals and interests.

II. EMOTIONS AND ETHICS: THE DECLINE OF PHINEAS GAGE

What is it that separates good character from bad? In Robert Louis Stevenson's masterpiece, *The Strange Case of Dr. Jekyll and Mr. Hyde*, a mysterious elixir transforms the beneficent Dr. Jekyll into the evil Mr. Hyde.²⁴ But, for Phineas Gage, a neuroscience enigma, the journey from civic virtue to

¹⁵ Leonard L. Riskin, *Awareness and Ethics in Dispute Resolution and Law: Why Mindfulness Tends to Foster Ethical Behavior*, 50 S. TEXAS L. REV. 493, 494, 498 (2009).

¹⁶ See *id.* at 498-501.

¹⁷ See Riskin, *supra* note 13; Riskin, *supra* note 15, at 501-03.

¹⁸ See Riskin, *supra* note 15, at 495.

¹⁹ See *id.* at 499.

²⁰ See *infra* Part II.

²¹ See *infra* Part III.

²² Riskin, *supra* note 15, at 501 ("The most challenging ethical choices arise when professionals face legitimate conflicting obligations. . .").

²³ See *infra* Part IV.

²⁴ ROBERT LOUIS STEVENSON, *THE STRANGE CASE OF DR. JEKYLL AND MR. HYDE* (Katherine B. Linehan ed., W.W. Norton & Co. 2002) (1886).

social ruin began, not with a vial, but with a blow to the head.²⁵ While a chemical brew unleashed Mr. Hyde's amoral energy, the story of Phineas Gage demonstrates that it is the chemicals in the brain, along with the feelings they induce, that help people make socially responsible decisions.

Gage's story, as retold by the neurologist Antonio Damasio, is a Kafkaesque tale of injury, metamorphosis, and decline.²⁶ In 1848, Phineas Gage, a well-respected and successful railroad foreman, mishandled explosives with terrible results.²⁷ An iron rod three feet in length and a quarter inch in diameter surged upward through his brain, resecting a crucial section of his frontal lobe.²⁸ Amazingly, Gage survived.²⁹ In fact, he was conscious immediately following the accident and calmly explained what had happened to the doctor who attended him.³⁰ In the doctor's words, Mr. Gage "related to me some of the circumstances, as he has since done; and I can safely say that neither at that time nor on any subsequent occasion, save once, did I consider him to be other than perfectly rational."³¹

Stunningly, Gage's mental faculties emerged intact. His "attention, perception, memory, language, [and] intelligence" all functioned normally.³² But, his personality had radically changed for the worse. He went from being a man of "temperate habits" and "considerable energy of character"³³ to being "fitful, irreverent, indulging at times in the grossest profanity . . . manifesting but little deference for his fellows, impatient of restraint or advice . . ." ³⁴ Too impulsive and erratic to hold a job or sustain relationships, Gage drifted from one position to the next, at one point earning money by displaying his wounds at the circus freak-show.³⁵ He was never able to stay in one place, save money, or set down roots.³⁶ Perennially on the outs, Gage limped along on society's fringes.³⁷ He died, alone and unnoticed, at the age of thirty-eight.³⁸

What happened to Gage? According to Damasio, the unfortunate railway man suffered damage to the ventromedial region of the brain's frontal lobe where emotion is produced and transmitted.³⁹ After his injury, Gage could think but no longer feel. And, because he had lost emotion's guidance func-

²⁵ See generally JOHN FLEISHMAN, PHINEAS GAGE: A GRUESOME BUT TRUE STORY ABOUT BRAIN SCIENCE 1-10 (Houghton Mifflin 2002) (recounting the story of Phineas Gage, a railway foreman who suffered a brain injury that left him physically intact but morally challenged).

²⁶ See ANTONIO R. DAMASIO, DESCARTES' ERROR: EMOTION, REASON, AND THE HUMAN BRAIN (1994).

²⁷ *Id.* at 3-4.

²⁸ *Id.* at 4, 32.

²⁹ *Id.* at 5.

³⁰ *Id.* at 5-6.

³¹ *Id.* at 6.

³² *Id.* at 11.

³³ *Id.* at 8.

³⁴ *Id.*

³⁵ *Id.* at 8-9.

³⁶ *Id.* at 8-10.

³⁷ *Id.* at 10.

³⁸ *Id.*

³⁹ *Id.* at 32.

tion, he could no longer anticipate and plan for the future, empathize with others, or harness his impulses to fit with existing social mores.⁴⁰

Damasio studied twelve patients with similar damage to their frontal lobes and each evinced the same set of deficits.⁴¹ These patients were cognitively functional and scored well on standardized psychological and neuropsychological tests.⁴² However, they could neither produce or access their emotions nor imagine the emotions of others.⁴³ Sadly, this defect in emotion production led to an inability to make appropriate decisions in the personal and social domains.⁴⁴ Typically, these patients could not hold a job, maintain intimate relationships, or manage the tasks essential for independent living.⁴⁵

To be sure, not all of these difficulties map directly to the ethical domain. Failure to perform at work, to select or emotionally satisfy a spouse, or to maintain an income do not alone establish ethical dysfunction. Nonetheless, evidence exists that damage to those sections of the brain responsible for emotion- production impairs decision-making in ways likely to lead to ethically questionable behavior.⁴⁶

Let us look at the gambling experiments conducted by Damasio and colleagues.⁴⁷ In these studies, clinical subjects were presented with four decks of cards.⁴⁸ The subjects were given no information about these decks and only instructed that their object was to make as much and lose as little money as possible. Unbeknownst to the subjects, each deck presented vastly different opportunities for rewards and losses. The cards in decks A and B paid \$100. The cards in decks C and D paid only half that amount. However, wild cards were interspersed throughout decks A and B that required the subject to pay a hefty sum—up to \$1250. Decks C and D also contained cards that required a pay-out, but those fines were minimal—capping out at \$250. Thus, decks A and B offered larger incremental gains, but presented enhanced risks of signifi-

⁴⁰ *Id.* at 10.

⁴¹ *Id.* at 53-54.

⁴² *Id.* at 54.

⁴³ *Id.* at 207-212.

⁴⁴ *Id.* at 54.

⁴⁵ *Id.* at 55-58; Antonio R. Damasio, Daniel Tranel & Hanna C. Damasio, *Somatic Markers and the Guidance of Behavior: Theory and Preliminary Testing*, in *FRONTAL LOBE FUNCTION AND DYSFUNCTION* 217-18 (Harvey S. Levin et al. eds., 1991); Antonio R. Damasio, *The Somatic Marker Hypothesis and the Possible Functions of the Prefrontal Cortex*, 351 *PHIL. TRANSACTIONS ROYAL SOC'Y B* 1413, 1413 (1996).

⁴⁶ See generally DAMASIO, *supra* note 26, at 48-49; Steven W. Anderson, Antoine Bechara, Hanna Damasio, Daniel Tranel & Antonio R. Damasio, *Impairment of Social and Moral Behavior Related to Early Damage in Human Prefrontal Cortex*, 2 *NATURE NEUROSCIENCE* 1032-33, 1035 (1999); Antoine Bechara, Hanna Damasio, Antonio R. Damasio & Gregory P. Lee, *Different Contributions of the Human Amygdala and Ventromedial Prefrontal Cortex to Decision-Making*, 19 *J. NEUROSCIENCE* 5473, 5480-81 (1999).

⁴⁷ See DAMASIO, *supra* note 26, at 212-19; see generally Antoine Bechara, Hanna Damasio, Daniel Tranel & Antonio R. Damasio, *Deciding Advantageously Before Knowing the Advantageous Strategy*, 275 *SCIENCE* 1293 (1997) [hereinafter Bechara et al., *Deciding Advantageously*]; Antoine Bechara, Antonio R. Damasio, Hanna Damasio & Steven W. Anderson, *Insensitivity to Future Consequences Following Damage to Human Prefrontal Cortex*, 50 *COGNITION* 7 (1994) [hereinafter Bechara et al., *Insensitivity to Future Consequences*].

⁴⁸ Bechara et al., *Insensitivity to Future Consequences*, *supra* note 47, at 8-10.

cant loss. Decks C and D presented lower incremental gains, but exacted less of a penalty when the pay-out cards were drawn.

Damasio's lab ran the gambling test on individuals with normal brains, as well as subjects with frontal-lobe damage.⁴⁹ Responses among the non-damaged control group diverged dramatically from those of the brain-damaged subjects. The control group began by taking cards from all four decks, but over time (within the first thirty of a hundred moves) they gravitated to decks C and D. They developed a sense that decks A and B were "scarier" than decks C and D, and this sense of risk informed their choices. Although decks C and D offered smaller rewards, the control group assumed that avoiding the "scary decks" would yield a higher cumulative amount in the long run.

In contrast, the individuals with frontal-lobe damage, after initial exploratory moves, favored decks A and B over the safer C and D decks.⁵⁰ Although they initially gained more from the higher pay-outs (\$100 as opposed to \$50), the higher penalties eventually led them into penury. Interestingly, these self-sabotaging choices were not ill-informed. Patients with frontal-lobe damage *understood* what was occurring; they *understood* that their choice of the cards from decks A and B produced disastrous penalties that undermined the potential for gain. Nonetheless, they continued to make the same self-defeating choice.

Why? According to Damasio, the group with frontal-lobe damage made their card choices without the navigational assistance that emotions typically provide.⁵¹ In his view, patients with normal brain function developed hunches regarding the "goodness" or "badness" of each of the decks.⁵² These "hunches" are comprised of unconscious, rough-grained assessments of the reward/punishment ratio embodied in each deck, combined with bodily reactions, or "somatic states," that serve to mark the assessment and link it to the appropriate stimulus.⁵³ Thus, when a non-brain-damaged subject begins to reach for the A or B decks, she feels a sense of threat (at a preconscious level), which stays her hand. However, individuals with frontal-lobe damage do not have the neurological machinery required to generate somatic markers.⁵⁴ As they do not have a mechanism linking previous experiences with emotion, there are no feelings to guide future choices. Nothing signals to frontal-lobe damaged individuals, "Don't pick from deck A—you were hit with a painful pen-

⁴⁹ *Id.* at 10-12.

⁵⁰ *Id.* at 12-14.

⁵¹ *Id.* at 13.

⁵² DAMASIO, *supra* note 26, at 219. For general discussion of the concept of "goodness" and "badness," see *id.* at 111, 186.

⁵³ *Id.* at 221. Damasio's own explanation of the "somatosensory system" is especially useful to understanding the concept of "somatic states," as explored throughout DESCARTES' ERROR and the remainder of this paper: "from the Greek root *soma*, for body; the somatosensory system is responsible for both the external senses of touch, temperature, pain, and the internal senses of joint position, visceral state, and pain. . . ." *Id.* at 65. "Note that whenever I use the term somatic or somatosensory I have in mind the soma, or body, in the general sense, and I refer to all types of body sensation including visceral sensations." *Id.*

⁵⁴ See generally Daniel Tranel & Hanna Damasio, *Neuroanatomical Correlates of Electrodermal Skin Conductance Responses*, 31 *PSYCHOPHYSIOLOGY* 427, 435 (1994) (noting that evidence suggests the ventromedial frontal region may play a significant role in the modulation of electrodermal responses to psychologically-derived stimuli).

alty the last time you did that!” Accordingly, these individuals pick from the risk-laden decks again and again and again.

An inability to produce and experience emotional states leads to what Damasio terms “myopia for the future.”⁵⁵ That is, individuals set adrift from their emotions are more likely to pursue outcomes that offer immediate high rewards and bypass avenues for longer-term gain.⁵⁶ Of course, this is exactly the sort of mind-set that leads to ethical (and criminal) missteps.⁵⁷ A mediator who overcharges a client or breaches confidentiality to close a deal essentially sacrifices her long-term reputation for immediate gain. It is thus unsurprising that individuals perennially unable to conform to social norms often suffer from a deficit in emotional processing.⁵⁸

Psychopaths, a population subset characterized by nihilistic disregard for rules and boundaries, behave much like the frontal-lobe damaged patients when exposed to the gambling protocol. They continue to pick from the alluring but ruinous A and B decks.⁵⁹ Moreover, their self-destructive behavior stems not from a defect in rationality or intelligence but from a deficit in emotional processing. When exposed to noxious stimuli (loud noises or visually disturbing pictures), individuals diagnosed with psychopathology demonstrate muted emotional response when compared to normal subjects.⁶⁰ Clinicians believe this problem is directly related to their inability to conform their behavior to society’s dictates.⁶¹

Guilt, shame, empathy, and compassion have all been termed “moral emotions.”⁶² Guilt and shame remind us that the intoxication of misbehavior is the prologue to a hangover of remorse. Similarly, empathy and compassion constrain self-interest, helping us see the world through other’s eyes. Without these moral emotions, we would cooperate less and defect more. Thus, as

⁵⁵ DAMASIO, *supra* note 26, at 217-18.

⁵⁶ *Id.* at 217.

⁵⁷ See Stephen D. Hart & Rebecca J. Dempster, *Impulsivity and Psychopathy, in IMPULSIVITY: THEORY, ASSESSMENT, AND TREATMENT* 212, 227 (Christopher D. Webster & Margaret A. Jackson eds., 1997) (speculating that, because of some “neurophysiologically based deficit in emotion or attention, psychopaths are more likely to consider the possibility of antisocial behavior and to evaluate such behavior as potentially rewarding” and that “psychopaths are less likely to inhibit antisocial cognitions”); see generally Sabine C. Herpertz et al., *Emotion in Criminal Offenders with Psychopathy and Borderline Personality Disorder*, 58 ARCHIVES GEN. PSYCHIATRY 737-45 (2001) (noting that criminal offenders with psychopathic features demonstrate emotional hyporesponsiveness to both appealing and aversive stimuli).

⁵⁸ Hart & Dempster, *supra* note 57, at 221-22 (reviewing etiological models of psychopathy that posit a neurobiological deficit as the cause of psychopaths’ “impulsive”—that is disinhibited, irresponsible, and antisocial—behavior”).

⁵⁹ DAMASIO, *supra* note 26, at 214.

⁶⁰ Jack van Honk et al., *Defective Somatic Markers in Sub-Clinical Psychopathy*, 13 NEUROREPORT 1025, 1026 (2002).

⁶¹ *Id.*

⁶² Jonathan Haidt, *The Moral Emotions, in HANDBOOK OF AFFECTIVE SCIENCES* 852, 852-53 (Richard J. Davidson et al. eds., 2003).

rational self-maximizers, devoid of emotional commitments, we would create a chilly, trustless world.⁶³

But emotion's role as ethical bullhorn is imperfect at best. To say that emotion is a necessary precondition for ethical decision-making is not to admit its sufficiency. While our emotions may be easily manipulated, it is not clear that even consistent mindful practice can help us master the unconscious, emotionally-driven assessments that mark our thought processes.

III. STRANGERS FROM OURSELVES: EMOTIONS THAT ELUDE MINDFUL CONTEMPLATION

Professor Riskin urges mindful practice because it can reduce "automatic, habitual ways of thinking, feeling and behaving" that inhibit collaborative, pie-expanding negotiation, as well as ethical decision-making.⁶⁴ But what if many of our reactions to the world consist of automatic thoughts that are inaccessible to our conscious selves?

A. *The Adaptive Unconscious and Confabulation*

In his text, *Strangers to Ourselves: Discovering the Adaptive Unconscious*,⁶⁵ psychologist Timothy Wilson presents persuasive data of a dual cognitive and affective processing system. He argues that our behaviors are directed by two distinct selves—one conscious, the other unconscious.⁶⁶ On one hand, our conscious self acts deliberately and in a controlled fashion on explicit motives, emotions, feelings, and desires that are known to us.⁶⁷ Our adaptive unconscious, on the other hand, acts in an automatic way on the implicit motives and emotions that Wilson believes remain forever inaccessible to our conscious selves.⁶⁸ Were Wilson (and many other psychologists) to read Riskin's work, he would likely agree that mindfulness can make us more attentive to our explicit emotions. However, he would caution that mindfulness is powerless to influence or even penetrate the inner workings of the adaptive unconscious, a structure that powerfully molds our most basic responses.

If our adaptive unconscious played an inconsequential role in our overall behavior, then Wilson's theory would hardly undermine the reasoning behind mindful practice. However, current understandings of the division of labor between our conscious and unconscious selves allocate considerable heft to the unconscious nether-regions of our psyche. Some researchers view the conscious self more as an "after-the-fact check and balancer" rather than a primary author of action.⁶⁹ Psychologists have alternately used the metaphors of presi-

⁶³ See generally ROBERT H. FRANK, *PASSIONS WITHIN REASON: THE STRATEGIC ROLE OF THE EMOTIONS* (1988) (explaining how emotions cause people to act contrary to their own self-interest).

⁶⁴ Riskin, *supra* note 13, at 325.

⁶⁵ TIMOTHY D. WILSON, *STRANGERS TO OURSELVES: DISCOVERING THE ADAPTIVE UNCONSCIOUS* (2002).

⁶⁶ *Id.* at 22-23, 49-64.

⁶⁷ *Id.* at 44-49.

⁶⁸ *Id.* at 23-24.

⁶⁹ *Id.* at 49-50.

dent and press secretary or hard-working federal government staff and figure-head executive to capture the extensive role the adaptive unconscious plays in maintaining functionality and driving decision-making.⁷⁰ Although our conscious self commands slow and effortful deliberations, our adaptive unconscious enjoys complete reign over the automatic, instantaneous judgments that shape our course.⁷¹ As Wilson suggests, the fact that many of our reactions to the world occur out of our “mind’s eye” means we often fail to comprehend what we are feeling or the reasons we feel what we do.

For example, take the divergence between self-report questionnaires that measure explicit, self-attributed motives and the information derived from psychological tests that measure unconscious needs and drives. The Thematic Appreciation Test (TAT) is designed to measure implicit drives for affiliation, power, and achievement.⁷² Subjects are shown pictures and asked to construct stories about the pictures, which are then coded based on the levels of need for power, affiliation, or achievement they reveal.⁷³ Consistently, in study after study, TAT results differ significantly from self-reports and serve as a better predictor of automatic behavior than the information subjects provide in questionnaires.⁷⁴ For example, TAT results for affiliation correlated and predicted whether the subject would be talking with another person when a researcher paged him or her at random intervals over several days.⁷⁵ Thus, even if an individual self-reported low-level needs for affiliation, a high TAT score for affiliation likely resulted in researchers finding the individual engaged with others, rather than in a solitary state.⁷⁶ According to researchers, both the TAT and self-report measures are accurate in their own sphere; each measures needs and motives, the first implicit, the second explicit.⁷⁷ However, the drives measured by the TAT will not achieve expression in a questionnaire because they reside outside the realm of conscious awareness.

Our lack of clarity regarding how and why we experience certain emotions extends to our most basic forms of excitation. In a series of experiments designed to test individual arousal to the opposite sex under both calm and anxious conditions, an attractive female aide approached young male subjects,

⁷⁰ *Id.* at 46-47; see also Robert Kurzban & C. Athena Aktipis, *Modularity and the Social Mind: Are Psychologists Too Self-ish?*, 11 *PERSONALITY & SOC. PSYCHOL. REV.* 131, 135-36 (2007) (referencing metaphors of a press secretary and president for the “self,” but also describing specialized modular cognitive subsystems as functioning akin to a “Machiavellian spin-doctor,” selectively packaging information for social advantage).

⁷¹ See WILSON, *supra* note 65, at 73-89.

⁷² *Id.* at 82-83, 91, 173-74.

⁷³ *Id.* at 82.

⁷⁴ *Id.* at 82-83; see generally David C. McClelland, Richard Koestner & Joel Weinberger, *How Do Self-Attributed and Implicit Motives Differ?*, in *MOTIVATION AND PERSONALITY: HANDBOOK OF THEMATIC CONTENT ANALYSIS* 49, 51-52 (Charles P. Smith ed., 1992).

⁷⁵ See WILSON, *supra* note 65, at 83.

⁷⁶ See McClelland et al., *supra* note 74, at 53-54.

⁷⁷ See PHEBE CRAMER, *STORYTELLING, NARRATIVE, AND THE THEMATIC APPERCEPTION TEST* 265 (paperback ed. 2004) (“Implicit motives are derived from more primitive, biologically based needs, are intrinsically associated with affect, and are most easily expressed in free-response situations. . . . Self-attributed motives are rooted in self-conceptions, which are largely based on learning, and are best assessed using self-report techniques.”).

ostensibly to fill out a questionnaire about creativity.⁷⁸ After the subject completed the questionnaire, the aide offered to discuss the questionnaire in greater detail and gave the subject her phone number.⁷⁹ As a sign of how attracted the subjects were to the young woman, the researchers recorded how many of the subjects called her and asked for a date.⁸⁰ Although the same woman participated in each encounter, the encounters took place in very different contexts. Half the men were approached while they were standing on a flimsy footbridge as it swung precariously between two deep gorges.⁸¹ The other half were questioned at a bench after crossing the bridge.⁸²

The aide changed nothing about her appearance or presentation when talking to the men on the footbridge as opposed to those at the bench.⁸³ Yet only thirty percent of the men on the bench called her for a date compared to sixty-five percent of the men on the footbridge.⁸⁴ Why were the “footbridge” men attracted to the aide at such higher rates than the “bench” men? Why did the geography of the encounter play such a significant role?

The researchers surmise that the men on the footbridge were physically aroused as a result of their peril. Their precarious physical location generated physical responses: their hearts raced, their pulses pounded, and they sweated.⁸⁵ Although one would think that the men would understand why they were feeling so excited, the men seemed confused on this point. The footbridge men knew they were in a heightened state, but they misattributed the source.⁸⁶

In many situations, we support our misattribution with carefully constructed reasons. Not only do we misperceive the source of our feelings, but we clothe our error in a neat web of confabulation.⁸⁷ Researchers have illustrated this confabulation in the celebrated “pantyhose” experiment: Study authors set up a display booth with pairs of panty hose arrayed side by side and asked subjects to examine the panty hose and then rank them according to preference.⁸⁸ Earlier experiments had demonstrated that positioning of consumer goods had a noticeable effect on preference, with goods furthest to the right garnering higher ratings.⁸⁹ Indeed, pantyhose proved no exception: only

⁷⁸ See WILSON, *supra* note 65, at 100-02; see generally Donald G. Dutton & Arthur P. Aron, *Some Evidence for Heightened Sexual Attraction Under Conditions of High Anxiety*, 30 J. PERSONALITY & SOC. PSYCHOL. 510-17 (1974).

⁷⁹ See WILSON, *supra* note 65, at 101.

⁸⁰ *Id.*

⁸¹ *Id.*

⁸² *Id.*

⁸³ *Id.*

⁸⁴ *Id.* at 102.

⁸⁵ *Id.* at 101.

⁸⁶ *Id.* 101-02; see generally Dolf Zillmann, *Attribution and Misattribution of Excitatory Reactions*, in 2 NEW DIRECTIONS IN ATTRIBUTION RESEARCH 335 (John H. Harvey et al. eds., 1978) (examining the principal theories of emotion as they relate to self-perception of excitatory reactions).

⁸⁷ See WILSON, *supra* note 65, at 104; Richard E. Nisbett & Timothy DeCamp Wilson, *Telling More Than We Can Know: Verbal Reports on Mental Processes*, 84 PSYCHOL. REV. 231, 251-53 (1977).

⁸⁸ WILSON, *supra* note 65, at 102-04.

⁸⁹ *Id.* at 102.

twelve-percent of participants preferred Brand A, the pantyhose laid out furthest to the left; seventeen-percent preferred Brand B, the next pair over; Brand C attracted thirty-one-percent; and, the winner, Brand D, stationed on the far right of the display, was preferred by forty-percent of research subjects.⁹⁰

When questioned, the subjects offered various reasons for their preferences. Some pointed to texture, others to the elasticity or durability of their preferred brand.⁹¹ When asked if positioning had anything to do with their choice, the vast majority said no.⁹² Instead, they maintained that their preferences were based entirely on the garments' look, feel, and anticipated functionality.⁹³ Yet they could not have made decisions based on those attributes because the panty hose were all the same.⁹⁴ Although the research subjects formed preferences, they were oblivious to the reasons for those preferences. Accordingly, they made up reasons after the fact.

While most people assume they do not confabulate, research indicates that the practice is common.⁹⁵ Responses that arise in the adaptive unconscious are not subject to interrogation by our conscious selves.⁹⁶ Therefore, when put on the spot to explain these responses, we fabricate and are unaware that we are doing so.⁹⁷

These studies indicate that there are limits to our ability to understand and control some of our most ingrained responses. Professor Riskin urges mindful practice as a way to mitigate "automatic, habitual ways of thinking, feeling, and behavior,"⁹⁸ but ample evidence suggests that much of that automatic behavior is evolutionarily hard-wired and not susceptible to conscious control.⁹⁹ With practice, mindfulness may help us avoid becoming unduly angry or upset when our negotiating partners do not accord us sufficient respect or attention. Nonetheless, mindfulness likely cannot give us insight into the prejudices and stereotypical judgments that we make at an implicit, unconscious level. This fact affects the degree to which mindfulness can help us hew to a higher ethical path.

B. Prejudice at the Unconscious Level

Imagine, for a moment, that Mike Mediator is presiding over a dispute between a homeowner and a condominium board. Mike is ethically required to

⁹⁰ *Id.* at 102-03.

⁹¹ *Id.* at 103.

⁹² *Id.* One person, who had just learned about positional effects, said yes. *Id.*

⁹³ *Id.*

⁹⁴ *Id.*

⁹⁵ See WILLIAM HIRSTEIN, *BRAIN FICTION: SELF-DECEPTION AND THE RIDDLE OF CONFABULATION* 13-14 (2005).

⁹⁶ *Id.* at 4.

⁹⁷ *Id.*

⁹⁸ Riskin, *supra* note 13, at 325.

⁹⁹ See *supra* notes 83-85 and accompanying text; Arne Öhman & Stefan Wiens, *On the Automaticity of Autonomic Responses in Emotion: An Evolutionary Perspective*, in *HANDBOOK OF AFFECTIVE SCIENCES*, *supra* note 62, at 256, 256-57; see generally Sachiko Kinoshita & Marie Peek-O'Leary, *Two Bases of the Compatibility Effect in the Implicit Association Test (IAT)*, 59 Q.J. EXPERIMENTAL PSYCHOL. 2102 (2006) (reporting on four experiments that investigate the mechanisms underlying the Implicit Association Test).

mediate fairly and impartially. Virtually every mediation code enjoins Mike from favoring either the homeowner or the board representatives.¹⁰⁰ Now imagine that our mediator and the condominium board representatives are white, while the homeowner is black. Studies examining individuals' implicit racial associations suggest that this racial configuration might strain Mike's ability to function impartially.

Research using a test called the IAT (implicit association test) demonstrates that even the most self-professed egalitarian may trade in stereotypes that subtly influence behavior.¹⁰¹ Although Mike may believe that he is acting impartially, he may signal non-neutral support for the board with selective eye-contact and body-language. In fact, Mike may be committed to practicing impartially and his deliberate behavior may demonstrate that commitment; he may think very hard about his introduction, subsequent interventions, and pay close attention to the evenhandedness of his comments. Still, controlled studies using the IAT reveal that professed attitudes do not always predict more subtle, non-verbal behavior, particularly when it comes to race and ethnicity.¹⁰²

Like the TAT, the IAT studies unconscious forms of thoughts and feelings. Specifically, it requires test subjects to sort concepts and attributes into categories and then evaluates the underlying associations.¹⁰³ The test assumes that subjects will more quickly sort items or concepts from categories that the research subject perceives "go together."¹⁰⁴ That is, research subjects will more easily, and therefore more rapidly, associate positive words (such as "happy") with items they view positively (such as flowers) and associate negative words (such as "ugly") with items they view negatively (such as poisonous insects).¹⁰⁵ Creating categories that link "happy" with "black widow spider" will, for all but the most ardent arachnophiles, challenge entrenched conceptual structures, require greater cognitive functioning, and generate a longer response time.

IATs studying race showed that many subjects harbored negative unconscious associations with blacks, despite conscious declarations of non-bias. In one study, researchers measured explicit prejudice among undergraduates by a twenty-item set of questions.¹⁰⁶ Next, they measured implicit prejudice by

¹⁰⁰ See, e.g., MODEL STANDARDS OF CONDUCT FOR MEDIATORS Standard II (2005) ("A mediator shall conduct a mediation in an impartial manner . . ."), available at <http://www.abanet.org/dispute/news/modelstandardsconductformediatorsfirolos.pdf> (adopted by the American Arbitration Association, American Bar Association, and Association for Conflict Resolution).

¹⁰¹ See Kinoshita & Peek-O'Leary, *supra* note 99, at 2103.

¹⁰² See Anthony G. Greenwald et al., *Understanding and Using the Implicit Association Test: III. Meta-Analysis of Predictive Validity*, 97 J. PERSONALITY & SOC. PSYCHOL. 17, 18 (2009).

¹⁰³ See UNDERSTANDING PREJUDICE AND DISCRIMINATION 22 (Scott Plous ed., 2002); Anthony G. Greenwald, Debbie E. McGhee & Jordan L. K. Schwartz, *Measuring Individual Differences in Implicit Cognition: The Implicit Association Test*, 74 J. PERSONALITY & SOC. PSYCHOL. 1464, 1464 (1998).

¹⁰⁴ Greenwald et al., *supra* note 103, at 1464.

¹⁰⁵ See Sally Lehrman, *The Implicit Prejudice*, SCI. AM., June 2006, at 32, 32.

¹⁰⁶ See John F. Dovidio, Kerry Kawakami, Craig Johnson, Brenda Johnson & Adaiah Howard, *On the Nature of Prejudice: Automatic and Controlled Processes*, 33 J. EXPERIMENTAL SOC. PSYCHOL. 510, 514 (1997).

flashing black and white faces on a computer screen as a subliminal prime before asking the students to make a decision about the word that followed.¹⁰⁷ The primary measure of implicit prejudice was “the degree to which participants responded faster to negative words following the Black prime than following the White prime, combined with the degree to which participants responded faster to positive words following the White prime than following the Black prime.”¹⁰⁸

In a subsequent study done by Dovidio, Kawakami, and Gaertner, Caucasian research subjects were evaluated on how their implicit racial associations and explicit racial attitudes relate to behaviors and impressions in interracial interactions. Following a self-report questionnaire and the race-oriented IAT, a researcher told the subjects that they were going to engage in an “acquaintance” study in which they would make the acquaintance of two “confederates” and spend three minutes with each confederate discussing a topic prescreened for race and gender neutrality.¹⁰⁹ The confederates were drawn from the same pool of undergraduates.¹¹⁰ In one discussion, the confederate was white, in the other, black.¹¹¹ The conversations and non-verbal interactions were recorded on camera and tape.¹¹² After each session, the research subject and the confederate were both asked to fill out questionnaires describing, on a seven-point scale, their impressions of how they and the confederate behaved in terms of friendliness.¹¹³ Additionally, judges viewed audio and videotaped recordings of the sessions and rated participants’ verbal and non-verbal behaviors based on the same criteria of friendliness and warmth.¹¹⁴ Judges used the level of verbal and non-verbal friendliness toward the white confederate as the baseline; they assessed racial bias based on differences in the ratings participants received regarding friendliness (verbal and nonverbal) toward the black confederate as compared with the white.¹¹⁵

The study yielded several conclusions. First, analysis of response times recorded by the IAT revealed implicit bias against blacks by whites.¹¹⁶ Second, explicit prejudice, as measured by the self-report questionnaire, related to both observer ratings of verbal friendliness and participant self-perceptions, but these measures did not correlate with the IAT results.¹¹⁷ Instead, the IAT correlated with observer ratings of non-verbal behavior and confederate perceptions of bias.¹¹⁸ Thus, a research subject who claimed to hold race-neutral values and was friendly to her black confederate in conversation might none-

¹⁰⁷ *Id.* at 513.

¹⁰⁸ *Id.* at 517.

¹⁰⁹ See John F. Dovidio, Kerry Kawakami & Samuel L. Gaertner, *Implicit and Explicit Prejudice and Interracial Interaction*, 82 *J. PERSONALITY & SOC. PSYCHOL.* 62, 62-68 (2002).

¹¹⁰ *Id.*

¹¹¹ *Id.*

¹¹² *Id.*

¹¹³ *Id.*

¹¹⁴ *Id.*

¹¹⁵ *Id.* at 65.

¹¹⁶ *Id.*

¹¹⁷ *Id.*

¹¹⁸ *Id.*

theless reveal significant levels of implicit bias and demonstrate nonverbal behaviors that observers and confederates rated as unfriendly. As the researchers expected, there was little correlation between participant's self-perceptions of bias and the perceptions of the confederates and observers.¹¹⁹ Rather, self-knowledge was limited to conscious attitudes and deliberative action.¹²⁰ Although the participants chose their words carefully, in-line with their ideas of themselves as civil libertarians committed to equality, the conscious self could not completely submerge the prejudices of its unconscious counterpart.¹²¹ Race-based aversion manifested itself in subtle, non-verbal gestures, which were apparent to observers and discussion-partners alike.¹²²

This and other IAT studies indicate that mindfulness—conscious attention to what we say and how we say it—might not be enough. Being ethical means being impartial, but intense self-consciousness cannot guard against the subtle workings of the unconscious. Mike Mediator may strive to treat the homeowner and the board representatives equally, but his nonverbal messaging might undermine those efforts. The way he uses his body, makes (or does not make) eye contact, laughs, and clears his voice may communicate solidarity or alienation in ways that elude his conscious awareness.

This is, of course, not an argument against mindfulness. It is merely a cautionary tale about what mindfulness can accomplish alone. While mindfulness may influence our conscious feelings, we may need other techniques to keep our subconscious biases in check. For example, IAT scholars recommend changes to individuals' working environments, increasing exposure to images and situations that contradict prevailing stereotypes.¹²³ Thus, we can help Mike to mediate impartially in an interracial environment by ensuring that his working environment presents ample images of successful, high-functioning black professionals. If Mike works at a private, for-profit mediation firm or as part of a court-connected program, it is important that he has black colleagues with whom he interacts on a daily basis in collegial, mutually-respectful ways. Unfortunately, the composition of the ADR field, to date, is notoriously, homogeneously white.¹²⁴ Changing the composition of the ADR workplace will likely promote impartial interracial mediation to a greater degree than will incorporating mindfulness practice into the routines of individual mediators.

Modifying the work place also holds greater promise than mindful emotional awareness in curbing our selfish and self-aggrandizing impulses. In his text, *Predictably Irrational*,¹²⁵ MIT economist Dan Ariely reveals our common tendency to cheat in small ways and advocates environmental measures as the

¹¹⁹ *Id.*

¹²⁰ *Id.*

¹²¹ *Id.*

¹²² *Id.* at 67.

¹²³ See Samuel L. Gaertner & John F. Dovidio, *Categorization, Recategorization, and Intergroup Bias*, in *ON THE NATURE OF PREJUDICE: FIFTY YEARS AFTER ALLPORT* 71, 77 (John F. Dovidio et al. eds., 2005).

¹²⁴ See F. Peter Phillips, *Diversity in ADR: More Difficult to Achieve than First Thought*, *DISP. RESOL. MAG.*, Spring 2004, at 14, 14.

¹²⁵ DAN ARIELY, *PREDICTABLY IRRATIONAL: THE HIDDEN FORCES THAT SHAPE OUR DECISIONS* (1st ed. 2008).

cure.¹²⁶ Ariely hypothesizes two types of dishonesty. The first is purposeful, self-aware, and utilitarian.¹²⁷ This type of dishonesty is exemplified by “a pair of crooks circling a gas station,” evaluating the risks and likely rewards of clearing out the gas station cash register and making a run for it based on a calculated decision-tree analysis.¹²⁸ The second type of dishonesty is stealthy by virtue of its seeming innocuousness. It is the kind

committed by people who generally consider themselves honest—the men and women . . . who have “borrowed” a pen from a conference site[,] taken an extra splash of soda from the soft drink dispenser, exaggerated the cost of their television on their property loss report, or falsely reported a meal with Aunt Enid as a business expense.¹²⁹

This second type could also include padding a client’s invoice with extra billable hours or giving up early on cases pre-paid at a daily rate.

In a series of experiments conducted at Harvard Business School, MIT, Princeton, UCLA, and Yale, Ariely discovered that students would cheat on test answers if they were sure that they would not be caught.¹³⁰ The good news, however, was that they kept their cheating within bounds. Students would change their incorrectly marked answer sheets on approximately four out of fifty questions, but they would not change all the answers to get a perfect score.¹³¹ They were modest in their deceit.¹³² From these results, Ariely surmised that most people value their integrity and honesty, but our honesty monitor (Ariely analogizes this entity to Freud’s super-ego) only flips on when we consider stepping seriously over the line.¹³³ Until we reach that point, the smaller stuff slips in under the radar.¹³⁴

Ariely’s solution to this small-scale, but pervasive dishonesty has nothing to do with the inward gaze. Rather, like the IAT researchers, he suggests changing our external environment. Noting that cheating declined when students were asked to recall the Ten Commandments, Ariely speculates that a procedural mechanism to encourage “contemplation of a moral benchmark” will facilitate greater probity.¹³⁵ He suggests that professionals should sign an oath pledging allegiance to the ethics code governing their field.¹³⁶ Although, theoretically, all professionals know their conduct should accord with prevailing ethical standards, there is something significant about the reaffirmation of allegiance immediately prior to the moment of temptation. Based on extensive field experimentation, Ariely concludes, “[I]f we are reminded of morality at the moment we are tempted, then we are much more likely to be honest.”¹³⁷ Although paying mindful attention to the twinges of guilt we may feel when we

¹²⁶ *Id.* at 196-97.

¹²⁷ *Id.*

¹²⁸ *Id.*

¹²⁹ *Id.* at 197.

¹³⁰ *Id.* at 198-201.

¹³¹ *Id.* at 200.

¹³² *Id.*

¹³³ *Id.* at 203.

¹³⁴ *Id.*

¹³⁵ *Id.* at 209.

¹³⁶ *See id.* at 212-15.

¹³⁷ *Id.* at 213.

unethically exploit others could prevent large-scale fraud, we likely will feel little guilt when contemplating smaller violations. These ethical missteps that elude conscious thought call for environmental changes, not internal introspection.

IV. MINDFULNESS, EMOTION, AND HARD CASES

Unconscious responses influence our decision-making in the easy ethical case, in which the ethical course is clear, but we are tempted to take shortcuts. But what about the ethically hard case where compelling moral imperatives push in opposite directions?¹³⁸ Professor Riskin argues that navigating the complex ethical case requires a calm and quiet mind and that mindful practices can help cultivate that quality.¹³⁹

Certainly, the ability to remain cool, calm, and collected when ethical challenges arise is useful, yet mindful practice can only assist us in attending to the thoughts and feelings that reach the level of consciousness.¹⁴⁰ What about the more subterranean workings of our mind? How do emotion and cognition work together to “solve” complex moral problems at a more unconscious level?

A. *Utilitarian and Deontological Decision-Making in Hard Cases: The Trolley Problem*

Research psychologist Joshua Greene has been examining this question through a series of experiments based on the “trolley problem” famously elaborated by Judith Jarvis Thomson.¹⁴¹ The problem goes something like this:

A runaway trolley is hurtling down the tracks. If it remains on course, it will kill five people who are unavoidably lodged there. You can divert the trolley with a switch so that it will go down another track, on which one person rests. That person will die, but the other five will be saved. Should you hit the switch?¹⁴²

As it turns out, most people say yes; pulling the switch in this instance, although agonizing, is morally justified.¹⁴³ A second version of the trolley problem, called the footbridge variation, tends to elicit a different response:

Imagine now that the trolley is heading down the track, ready to cut down the lives of the same five doomed individuals. This time, there is no switch, but a heavy weight

¹³⁸ What constitutes an “ethically hard” case is open for debate. In another text, I define ethically challenging dilemmas as those where important ethical commitments come into conflict. See *ETHICS IN MEDIATION* (Ellen Waldman ed.) (forthcoming 2010). Professor Riskin appears to adopt a similar definition when he writes, “The most challenging ethical choices arise when professionals face legitimate conflicting obligations: a lawyer whose duty to foster the interests of his client seems to conflict with the obligation of truthfulness to third parties; the mediator who experiences a conflict between the obligation of impartiality and the obligation toward a quality process.” Riskin, *supra* note 15, at 501.

¹³⁹ Riskin, *supra* note 13, at 334.

¹⁴⁰ See *supra* notes 99-108 and accompanying text.

¹⁴¹ Judith Jarvis Thomson, *Killing, Letting Die, and the Trolley Problem*, 59 *MONIST* 204, 206-08 (1976); see also PHILIPPA FOOT, *The Problem of Abortion and the Doctrine of Double Effect*, in *VIRTUES AND VICES* 19-23 (1978).

¹⁴² Joshua D. Greene, R. Brian Sommerville, Leigh E. Nystrom, John M. Darley & Johnathan D. Cohen, *An fMRI Investigation of Emotional Engagement in Moral Judgment*, 293 *SCIENCE* 2105, 2105 (2001).

¹⁴³ *Id.*

can stop the trolley. You just happen to be positioned on a footbridge above the trolley line, alongside a very, very fat man. (You are, conversely, quite slender, removing the possibility of productive self-sacrifice.) You can save the five people on the track if you fatally push the fat man off the bridge. Should you?¹⁴⁴

When faced with this dilemma, most people said no.¹⁴⁵

Greene points out that, as a matter of theoretical consistency, this makes no sense.¹⁴⁶ A purely utilitarian approach would justify both throwing the switch and pushing the fat man off the bridge. Each disposes of one person while saving five, a net gain of four, a conclusion of which utilitarians would approve.¹⁴⁷ On the other hand, a deontological perspective justifies neither the push nor the throw. Deontological ethics forbids treating people merely as means to others' ends; one cannot inflict harm on one individual, even if doing so creates correspondingly greater goods for others.¹⁴⁸ Thus, killing one to save five fails from a deontological standpoint, regardless of whether we are throwing switches or heaving people off bridges. While some deontologists have labored to justify the switch-throwing on the ground that the harm generated was foreseen, but unintended, those arguments are generally considered unsatisfying.¹⁴⁹

Greene brought a unique twist to trolleyology by subjecting individuals to an fMRI brain scan while they pondered dilemmas similar to the trolley/foot-bridge conundrums. What he found was that when individuals contemplated dilemmas that forced "up-close and personal" harms, similar to throwing the fat man to his death, central portions of the brain associated with emotion were active.¹⁵⁰ In contrast, when individuals contemplated more impersonal, indirect action, such as throwing a switch that would indirectly, but fatally, divert machinery, portions of the brain associated with stored memory and reflective cognition showed increased relative activation.¹⁵¹ Greene and others have read the data to support a dual processing theory of moral reasoning that links deon-

¹⁴⁴ This hypothetical rephrases Greene's version in *id.*

¹⁴⁵ *Id.*

¹⁴⁶ *Id.* at 2106.

¹⁴⁷ See *infra* notes 159-61 and accompanying text.

¹⁴⁸ See *infra* notes 154-58 and accompanying text.

¹⁴⁹ As Greene puts it, a Kantian might try to distinguish the two situations by noting that in the foot-bridge example, one is using the fat man purely as a train-stopper—that is, as merely a means as opposed to an end in himself. In the trolley-switching situation, the man being mowed down on the alternative track is not being used to stop the train. His death is an unfortunate by-product of a track switch that otherwise makes great sense. This answer, however, "runs into trouble with a variant of the trolley dilemma in which the track leading to the one person loops around to connect with the track leading to the five people." In that instance, most people still would throw the switch, even though the person on the looped around track is being used to stop the train, just as the fat man would be. See Greene et al., *supra* note 142, at 2106 ("[T]here is no set of consistent, readily accessible moral principles that captures people's intuitions concerning what behavior is or is not appropriate in these and similar cases."); see also SHELLY KAGAN, *NORMATIVE ETHICS* 100-105 (1998) (discussing the proposal that "instead of a constraint against *doing* harm, the deontologist should accept a constraint against *intending* harm").

¹⁵⁰ Greene et al., *supra* note 142, at 2107.

¹⁵¹ See *id.* at 2106 fig.1.

tological responses with emotional engagement and utilitarian reactions with more purely cognitive circuitry.¹⁵²

While this link has elicited criticism and inspired debate,¹⁵³ it has ramifications for ethical deliberation in the mediation field. If Greene is right, our deontological instincts are driven by emotional intuitions that can be overridden by “cooler” more deliberate cognition. That raises questions regarding our approach to the hard cases in mediation where our deontological commitments conflict with more utilitarian calculi. Therefore, let us first look at how deontological approaches, as opposed to utilitarian approaches, might manifest in the mediation setting and how tensions involving these two approaches might arise.

B. *Deontology and Utilitarianism in Mediation*

Kant’s deontological theory is concerned primarily with rights and duties.¹⁵⁴ According to Kant, certain actions are morally compelled and others morally forbidden, regardless of the consequences.¹⁵⁵ Individuals have perfect rights, such as the right to be treated as an autonomous being with inherent rights of dignity and personhood.¹⁵⁶ Additionally, individuals have perfect duties, such as the duty to tell the truth.¹⁵⁷ These rights and duties cannot be transgressed, even if their violation would generate a better set of outcomes than strict adherence.¹⁵⁸ For instance, if a group of Nazi soldiers came to your house and asked if you were shielding Jews from deportation and death, deontological thinking would compel you to tell the truth and yield your human contraband. In short, the duty to speak truthfully remains absolute, regardless of the tragic results.

Utilitarianism, on the other hand, is concerned primarily with consequences.¹⁵⁹ An action is morally compelled if it maximizes happiness, by creating the greatest good (happiness) for the greatest number.¹⁶⁰ Whether an action is morally right or wrong is not determined by a priori obligations but rather by consideration of future consequences.¹⁶¹ When considering the Nazis and the truth-telling dilemma, a utilitarian would consider whether disclosure would maximize happiness for all affected individuals, and, based on the certain harm facing the Jews, would conclude that lying was the morally preferred course of action.

¹⁵² See *id.* at 2107.

¹⁵³ Greg Miller, *Growing Pains for fMRI*, 302 SCIENCE 1412, 1413 (2008).

¹⁵⁴ See generally Onora O’Neill, *The Moral Perplexities of Famine and World Hunger*, in MATTERS OF LIFE AND DEATH 294, 319-20 (Tom Regan ed., 2d ed. 1986) (providing a basic explanation of Kantian ethics).

¹⁵⁵ *Id.* at 319.

¹⁵⁶ *Id.*

¹⁵⁷ W. D. ROSS, THE RIGHT AND THE GOOD 18 (1930).

¹⁵⁸ *Id.*

¹⁵⁹ See generally Onora O’Neill, *Kant and Utilitarianism Contrasted*, in MORALITY AND MORAL CONTROVERSIES: READINGS IN MORAL, SOCIAL, AND POLITICAL PHILOSOPHY (John Arthur ed., 2005).

¹⁶⁰ *Id.*

¹⁶¹ *Id.*

Ethical dilemmas in mediation present classic deontological and utilitarian tensions. Consider how a deontological and a utilitarian thinker would broach the following case:

Renee and Bill are divorcing after a twenty-year marriage. Renee has raised the couple's three children, but has never worked out of the home. Bill, a partner in a brokerage firm, has managed the couple's finances and investments. Renee has asked for a divorce, but feels guilty about breaking up the family. Bill feels angry and betrayed. He believes Renee has "met someone else" and has taken the position that Renee does not deserve half of the marital assets or child or spousal support. Renee does not want to fight and simply wants the divorce proceedings over and done with. She is inclined to settle for much less than what is her due under state law simply for the sake of obtaining closure. She feels genuinely sorry for breaking up the marriage and does not want to fight with Bill.

Regardless of ideological predilections, most mediators would likely engage in a set of standard interventions.¹⁶² First, they would encourage Renee to consult with an attorney regarding the concessions she is preparing to make. In addition, they would pose a series of questions designed to inquire into the pragmatic consequences of Renee's situation: What are Renee's current and likely future expenses? What sort of earning potential does she have or envision having in the future? How will she live if she accepts the minimal package Bill is offering? What are the children's needs? How will those needs be met? Why is she willing to settle for less than the local child support and spousal support norms would allow?

After recommending the retention of counsel and engaging in minimal reality-testing, a mediator may then move in a variety of directions. Existing ethical guidance leaves ample room to maneuver.¹⁶³

A mediator with a deontological bent would see his or her primary duty as respecting Renee's autonomy and promoting her ability to self-determine her post-divorce life. Kant emphasized every individual's right to be treated with dignity and respect. Every individual, he postulated, has the right to be treated as an end in herself, rather than as a vehicle for the satisfaction of others' needs. Thus, a mediator following Kantian thinking would, above all, try to avoid persuading or coercing Renee to reach an agreement that diverges from Renee's own preferences, regardless of how idiosyncratic or irrational those preferences might be.

By contrast, a mediator with a utilitarian bent would evaluate consequences by determining what Renee's waiver will mean for her and her children five, ten, or fifteen years hence. She or he would try to predict whether

¹⁶² See MODEL STANDARDS OF CONDUCT FOR MEDIATORS, *supra* note 100, at Standard VI.

¹⁶³ Standard I, for example, commands mediators to ground their mediation practice in the principle of self-determination by the parties, which requires voluntary, uncoerced, and informed decision-making. At the same time, Standard I explicitly relieves the mediator of any obligation to ensure informed consent. Thus, a mediator could emphasize the importance of obtaining sufficient information to make a knowledgeable settlement decision and push Rachel to gather more information about the consequences of waiving her financial entitlements. At the same time, the mediator could assume a more passive stance on the theory that Rachel is a fully autonomous adult who knows her own interests and likely has very good reasons for choosing to settle for less rather than fight for more. Both mediator strategies could be viewed as consistent with the plain language of Standard I.

allowing Renee to settle quickly now for a small amount will maximize happiness in the long run, or if it will lead to longer term poverty and remorse. This mediator would prioritize Renee's self-determination as secondary to facilitating an agreement that maximizes Bill's, Renee's, and the children's long-term best interests.

The deontological/utilitarian divide tracks the classic tension mediators experience between conflicting duties to be fair and impartial.¹⁶⁴ Do mediators have any ethical prerogative other than facilitating party self-determination and staying neutral? Should fairness, as measured by legal or social norms, as opposed to party preferences, play any role in the mediator's thinking about his or her ethical obligations? If the parties are ready to go along, is there any justification for intervening due to concerns about the settlement's long-term consequences? If the parties are prepared to say the terms are fair, on what authority may the mediator offer a different assessment?

Professor Greene concludes that our deontological impulses are evolutionary relics: useful at one time, but increasingly wrong-headed.¹⁶⁵ He holds that our modern world requires a more rigorously utilitarian approach that focuses on consequences.¹⁶⁶ I do not go so far as Greene. I would not jettison the impulses that keep the fat man safe from a utility-maximizing death. Still, Greene's suggestion that our deontological reflexes warrant examination may apply in the mediation arena. When we define our ethical responsibilities solely in terms of promoting self-determination, we may be succumbing to emotionally comfortable shibboleths without sufficient regard for the practical consequences. In particular, we should consider whether we are paying adequate attention to the capacities required to act autonomously and the effects such autonomous action will have on third parties.

I sometimes wonder whether the "autonomy reflex," the reflex that limits our ethical obligations to helping the parties before us meet their momentary preferences, reflects a sort of moral laziness—an unwillingness to confront the untoward possibilities that such a constrained understanding of mediation ethics implies. To be sure, unfettered consequentialism can lead to results as absurd as those generated by a rigid Kantian stance. It is for this reason that some philosophers suggest a more flexible balancing approach, one which tempers traditional Kantian concerns for autonomy with attention to consequences.¹⁶⁷ Our model standards provide a basis for this sort of balancing with their emphasis on facilitating party self-determination while presiding over a quality process.¹⁶⁸

¹⁶⁴ See generally Susan Nauss Exon, *How Can a Mediator Be Both Impartial and Fair?: Why Ethical Standards of Conduct Create Chaos for Mediators*, 2006 J. DISP. RESOL. 387 (arguing that mediators cannot simultaneously satisfy their ethical duties of neutrality and impartiality).

¹⁶⁵ See Joshua Greene, *The Secret Joke of Kant's Soul*, in 3 MORAL PSYCHOLOGY: THE NEUROSCIENCE OF MORALITY: EMOTION, DISEASE AND DEVELOPMENT (W. Sinnott-Armstrong eds., 2007).

¹⁶⁶ *Id.*

¹⁶⁷ See generally Ross, *supra* note 157.

¹⁶⁸ See MODEL STANDARDS OF CONDUCT FOR MEDIATORS, *supra* note 100, at Standards I and VI.

Some have read the mandate to preside over a quality process as merely reinforcing the primary mandate of self-determination.¹⁶⁹ According to this reading, the quality injunction is procedural only, oriented toward establishing those conditions required for fully autonomous decision-making.¹⁷⁰ It is also possible, however, to read Standard VI's quality requirement more substantively. One could read the Standard as requiring mediators to pay some attention to the fairness or equity of the outcome.¹⁷¹ In cases where power is distributed asymmetrically, parties may drift toward disturbingly one-sided outcomes. Additionally, the proposed deal may threaten the well-being of parties not at the table. Arguably, a quality process would, at the very least, avoid substantive outcomes that threaten long-term harm to both disputants and absent third-parties.¹⁷² In this way, mediation ethics would incorporate the duty of nonmaleficence,¹⁷³ which is so often salient in other professional codes.¹⁷⁴

V. CONCLUSION

The question of what constitutes "right action" in the mediation setting remains a subject of debate. Some commentators advance an "*autonomy uber alles*" philosophy,¹⁷⁵ while others temper their responses with concerns for

¹⁶⁹ Indeed, the language of the Standard itself and associated comments suggests that this reading may best correspond with the intentions of the drafters. Standard VI reads, "A mediator shall conduct a mediation in accordance with these Standards and in a manner that promotes diligence, timeliness, safety, presence of the appropriate participants, party participation, procedural fairness, party competency and mutual respect among all participants." *Id.* at Standard VI. Notably, Standard VI's definition of "quality" does not reference questions of substantive fairness at all.

¹⁷⁰ Standard VI's comments state that a quality process can only occur when: the mediator can commit sufficient time and attention to the mediation; the mediation takes place without undue delay; the appropriate stakeholders are present; the mediator can remain impartial; the parties have capacity to participate; mediator behavior is not influenced by the desire for a high settlement rate. *Id.* at Standard VI, cmt. a.; see also Nancy Welsh, *Making Deals in Court-Connected Mediation: What's Justice Got to Do with It?*, 79 WASH. U. L.Q. 787, 835 (describing quality requirement as relating to procedural, not substantive, fairness).

¹⁷¹ See generally Jonathan M. Hyman, *Swimming in the Deep End: Dealing with Justice in Mediation*, CARDOZO J. CONFLICT RESOL. 19, 39 (2004) (suggesting that mediators may profitably raise common-sense issues of fairness with parties).

¹⁷² Notably, the Ethics Standards for Family and Divorce Mediators do advise mediators to suspend or terminate the mediation process if the disputants "are about to enter into an agreement that the mediator reasonably believes to be unconscionable." MODEL STANDARDS OF PRACTICE FOR FAMILY AND DIVORCE MEDIATION Standard XI (Symposium on Standards of Practice 2000), available at <http://www.afccnet.org/pdfs/modelstandards.pdf>.

¹⁷³ Nonmaleficence derives from the ancient dictum *primum non nocere*, which, translated from Latin, means "Above all [or first] do no harm." TOM L. BEAUCHAMP & JAMES F. CHILDRESS, PRINCIPLES OF BIOMEDICAL ETHICS 189 (4th ed. 1994).

¹⁷⁴ See HIPPOCRATES, EPIDEMICS 165 (W. H. S. Jones, trans., Harvard Univ. Press 5th ed. 1962) (1923) ("Declare the past, diagnose the present, foretell the future; practise these acts. As to diseases, make a habit of two things— to help, or at least to do no harm.")

¹⁷⁵ See Linda Stamato, *Easier Said than Done: Resolving Ethical Dilemmas in Policy and Practice*, 1994 J. DISP. RESOL. 81, 82-83 (suggesting that mediators face ethical challenges when confronting power imbalances or inequitable outcomes, but arguing that mediators do not "value" cases and that party self-determination must remain paramount).

long-term consequences.¹⁷⁶ Professor Riskin's contribution urges mindful contemplation to enhance ethics.¹⁷⁷ Mindful contemplation, he suggests, will help mediators avoid self-aggrandizing misconduct and handle with equanimity the cognitive and emotional challenges posed by hard cases.¹⁷⁸

This Article suggests we heed Riskin's suggestions, albeit with modest expectations regarding what mindful practice, alone, can accomplish. It reviews inquiries into the workings of the adaptive unconscious and surveys efforts to uncover the neurological machinery that appears to power our diverse ethical responses. It discusses current theories regarding how our brain responds to philosophical dilemmas that threaten deontological and utilitarian commitments and speculates how mediation might incorporate those insights. Finally, it suggests that the role mediation plays in facilitating dispute outcomes imposes a duty on mediators to consider, not simply procedural, but substantive justice concerns. In this vein, adopting the principle of nonmaleficence would be a useful add-on to existing codes of conduct.

To date, theoretical examination in mediation has been largely geared toward exploring, critiquing, or defending the diverse models that populate the landscape. Scholars have parsed how these myriad approaches assume different goals, endorse different methods, and deploy different techniques. They have assumed that different mediation models imply different ethical commitments, but exactly how these commitments should be pursued or accommodated in any particular case remains relatively unexamined. We should be grateful to Professor Riskin for initiating this conversation on mindfulness, ethics, and their linkage in alternative dispute resolution. We can only hope that the conversation continues and deepens, as we bring our mindful attention to bear on these complex questions.

¹⁷⁶ See *MEDIATING ETHICALLY* (Ellen Waldman ed., forthcoming Dec. 2010) (manuscript on file with author).

¹⁷⁷ Riskin, *supra* note 13, at 334.

¹⁷⁸ *Id.*