RETHINKING THE ROLE OF INTENTIONAL WRONGDOING IN CRIMINAL LAW

113 CALIFORNIA LAW REVIEW _____ (forthcoming)

GREGORY ANTILL*

This Article employs recent philosophical advances in action theory and moral responsibility to critically examine the traditional purpose-knowledge-recklessness-negligence (PKRN) mens rea hierarchy of the Model Penal Code. It is a foundational assumption of the traditional mens rea hierarchy that the commission of intentional harm ought to be subject to greater criminal liability than actions which foreseeably result in risk of those same harms. This Article questions the soundness of that assumption. It argues that for many criminal offenses – particularly criminal homicide – a reluctant agent who purposefully causes harm to another person (even if deliberate and pre-meditated) will often nonetheless exhibit more concern for the well-being of their victims than a callous agent who acts recklessly, or even negligently, while indifferent to the harm they cause. The Article uses this critical re-thinking of the standard mens rea hierarchy to show how we might amend current homicide doctrine (and the PKRN mens rea regime more generally) to allow more criminal liability for non-intentional police homicides like Derek Chauvin's killing of George Floyd, relative to reluctant purposeful defendants.

As part of that argument, the Article identifies and articulates an especially important set of 'avoidance-commitments,' which are manifested in the case of reluctant purposeful agents but absent in the case of callous agents, and which speak in favor of diminished liability for many purposeful agents relative to their reckless or negligent counterparts. The Article shows how this novel analysis of culpability in terms of such avoidance commitments can be harnessed to develop an alternative set of mens rea classifications for criminal law. Such classifications could more closely track the underlying culpability of defendants than the current PKRN system, without forcing fact-finders to make problematic discretionary normative judgments about the quality of an agent's motives (a common problem bedeviling other recent prominent scholarly proposals for mens rea reform).

In so arguing, the Article highlights how the criminal law's current PKRN mens rea hierarchy, while seemingly ideologically neutral, in fact evinces a commitment on the part of the state toward punishing more severely a wide variety of crimes committed purposefully by defendants who were driven to commit those crimes out of poverty, abuse, and other forms of social marginalization. Whereas it has been used to avoid or reduce criminal liability by reckless or negligent defendants in positions of social power, such as police officers, white collar criminals, or landlords who cut corners on safety regulations to cut down costs, who commit non-intentional crimes of convenience while unwilling or unmotivated to take easily available options to avoid harming victims. Failure to be clear-eyed about such commitments creates a further barrier to recognizing the true moral magnitude of failures by police officers like Chauvin to recognize the humanity of those they police, and to designing a legal regime that fairly and effectively assigns criminal liability accordingly.

---

* Assistant Professor of Law, Pace University Elizabeth Haub School of Law; J.D. Yale Law School; Ph.D. UCLA Department of Philosophy. Earlier drafts of this material were presented to audiences at Yale Law School, Columbia Law School, Michigan Law School, the "Markelloquium!" Criminal Law Theory Colloquium, and the 2024 Legal Philosophy Workshop. Thanks to everyone who commented on those occasions. The paper also benefitted enormously from conversations and written comments from Larry Alexander, Michael Cahill, Luis C.deBaca, Raff Donelson, Ben Eidelson, Kim Ferzan, John Goldberg, Bernard Harcourt, Dan Kahan, Chris Lewis, Sandy Mason, Tracey Meares, Gabe Mendlow, Daniel Richman, Alex Sarch, Elizabeth Scott, Sarah Seo, Hannah Shaffer, Tom Tyler, and Gideon Yaffe.

TABLE OF CONTENTS

INTRODUCTION

On the evening of May 25th, 2020, Minneapolis police officer Derek Chauvin killed George Floyd by kneeling on the back of his neck for 9 minutes and 29 seconds,[1] ignoring Floyd's plea: "I can't breathe."[2] In his defense, Chauvin claimed that George Floyd's death was unintentional. His actions, Chauvin argued, were performed with the purpose of "assisting…in [Floyd's] arrest."[3] While not charged with the intentional killing of George Floyd, at trial the jury found that Floyd's death was the result of Chauvin's either recklessly disregarding the risk, or else negligently failing to even consider the risk altogether, that his chosen means for effecting Floyd's arrest would lead to Floyd's death.[4] Chauvin was convicted at trial of second degree unintentional murder, third degree murder, and second degree manslaughter, for acting in pursuit of

[1] State v. Chauvin, No. 27-CR-20-12646, 2021 WL 1559176 (Minn. Dist. Ct. Apr. 20, 2021) (opening statement at 3:00) ("You will learn what happened in that nine minutes and twenty-nine seconds… when Mr. Derek Chauvin was applying this excessive force to the body of Mr. George Floyd.").
[2] *Id.* (opening statement at 4:45).
[3] State v. Chauvin, No. 27-CR-20-12646, 2021 WL 1559176 (Minn. Dist. Ct. Apr. 20, 2021) (verdict as to Count III).
[4] *Id.* at 1. George Floyd had been arrested for allegedly using a counterfeit twenty dollar bill. *See* Nicholas Bogel-Burroughs & Will Wright, *Little Has Been Said About The $20 Bill That Brought Officers to the Scene*, N. Y. TIMES (Apr. 19, 2021) https://www.nytimes.com/2021/04/19/us/george-floyd-bill-counterfeit.html.

his goal in a manner which was "eminently dangerous to others…without regard for [George Floyd's] life."[5]

Despite the severity of the harm he committed, the fact that Chauvin did not intend George Floyd's death ensured that Chauvin was not convicted of the more legally serious crimes of intentional murder in the second degree, or murder in the first degree.[6] Indeed, putting aside his strict liability felony homicide conviction (itself an artifact of Minnesota's idiosyncratic merger rule),[7] Chauvin's convictions for reckless and negligent criminal homicide would involve, in almost any jurisdiction, among the least amount of criminal liability for having culpably caused such a harm.[8]

Compare cases of police perpetrated homicide like that of Chauvin's killing of George Floyd, with studies of female-perpetrated purposeful homicides like that of criminologists Angela Browne and Kirk Williams. Browne and Williams find that over half of all such homicides involve the killing of intimate partners, and that of these intimate-partner homicides, somewhere between 75% - 93% of the study's subjects report having been subject to physical or psychological abuse by the victim prior to the homicide.[9] In many cases, Browne and Williams found that the driving trigger for

---

[5] Minn. Stat. § 609.195 (2022) (murder in the third degree) ("Whoever, without intent to effect the death of any person, causes the death of another by perpetrating an act eminently dangerous to others and evincing a depraved mind, without regard for human life, is guilty of murder in the third degree and may be sentenced to imprisonment for not more than 25 years"); Minn. Stat. § 609.185 subdivision 2 (2022) (Unintentional murder in the second degree) ("Whoever does . . . the following is guilty of unintentional murder in the second degree and may be sentenced to imprisonment for not more than 40 years: (1) causes the death of a human being, without intent to effect the death of any person, while committing or attempting to commit a felony offense"); Minn. Stat. § 609.205 subdivision 1 (2022) (Manslaughter in The Second Degree) ("A person who causes the death of another by [the person's culpable negligence whereby the person creates an unreasonable risk, and consciously takes chances of causing death or great bodily harm to another] is guilty of manslaughter in the second degree and may be sentenced to imprisonment for not more than ten years or to payment of a fine of not more than $20,000, or both").

[6] Minn. Stat. § 609.185 (2022) (murder in the first degree) ("whoever does . . . the following is guilty of murder in the first degree and shall be sentenced to imprisonment for life: (1) causes the death of a human being with premeditation and with intent to effect the death of the person or of another"); Minn. Stat. § 609.185 subdivision 1 (2022) ("intentional murder in the second degree) ("Whoever does . . . the following is guilty of murder in the second degree and may be sentenced to imprisonment for not more than 40 years: (1) causes the death of a human being with intent to effect the death of that person or another, but without premeditation").

[7] *See* Gideon Yaffe, *The Lucky Legal Accident That Led to Derek Chauvin's Conviction*, THE HILL (May 1, 2021). The absence of a merger doctrine for assault and homicide in Minnesota law traces back to State v. Jackson, 346 N.W. 2d 634, 636 (Minn, 1984).

[8] In the federal criminal code, for example, negligent homicide is classified as "involuntary homicide" and has an upper sentencing limit of "not more than 8 years" imprisonment. 18 U.S.C. §1112. Intentional premeditated first-degree intentional murder, in contrast, "shall be punished" by a minimum of "imprisonment for life." 18 U.S.C. §1111.

[9] Angela Browne & Kirk K. Williams, *Exploring the Effect of Resource Availability and the Likelihood of Female-Perpetrated Homicides,* 23 L. & SOC. REV. 75, 77 (1989). More recent studies suggest that these numbers are likely still broadly accurate today. *See* e.g., Tamar Kraft-Stolar et al., *From Protection to Punishment: Post-Conviction Barriers to Justice for Domestic Violence Survivor-Defendants in New York State,* AVON GLOBAL CENTER FOR WOMEN AND JUSTICE AND DORTHEA S. CLARKE PROGRAM IN FEMINIST JURISPRUDENCE 2 (2011); Melissa Dichter, *Women's Experiences of Abuse as a Risk factor for Incarceration: A*

homicide was a situational context involving an absence of "legal [or] extra[-]legal resources," such as a hostile legal regime which refused to treat the existence of non-life-threatening abuse as grounds for unilateral divorce action.[10] While these reluctant purposeful actors did not see the death of their victim as their ultimate aim, they saw the deaths of their victims as the only means available for ending what they reported as an "overwhelming and entrapping life situation" from which they had no other avenues of escape.[11] Nonetheless, the fact that such homicides are purposeful – and often deliberate or pre-meditated – ensures that many such defendants continue face far more criminal liability than Chauvin or similar cases of police perpetrated reckless and negligent homicides.

While society has become increasingly aware of the injustice of the fact that police officers like Derek Chauvin, who kill suspects while effecting arrests, face far less criminal liability than the survivor-defendants described in Browne and Williams' study, the reason for this disparity remains, I will argue, under-appreciated. The reason for the disparity is not merely a hesitancy on the part of prosecutors to enforce existing criminal statutes against police officers and the existence of specialized shields from liability for police use of force.[12] Nor is it fully explained by the disparate application of available self-defense statutes to survivor-defendants, or the failure of many states to adopt or fully utilize legislative sentencing reforms for criminalized survivors of family violence, intimate partner violence, and sexual trafficking.[13] Rather, much of the difference in criminal liability between cases of police-perpetrated homicides like Derek Chauvin and the homicides committed by survivor-defendants reflects a deep and fundamental feature of both criminal and civil liability: a distinction in liability for harms which are intended and harms which are merely foreseen (or reasonably foreseeable) to some degree. It is a foundational assumption of the traditional mens rea hierarchy that the actions of an agent who causes some harm intentionally ought to be subject to more liability than the actions of an agent who causes the same harm unintentionally, through recklessness or negligence.

This assumption forms the basis for the enormously influential *purpose-knowledge-recklessness-negligence (PKRN)* mens rea hierarchy of the American Law Institute's Model Penal Code (MPC),[14] and for contemporary criminal law more

---

Research Update, NATIONAL ONLINE RESOURCE CENTER FOR VIOLENCE AGAINST WOMEN (2015); *Research Across the Walls,* SURVIVED AND PUNISHED (2011).

[10] Browne & Williams, *supra* note 9 at 78.

[11] *Id.*

[12] Though these are, of course, additional barriers which often prevent police-perpetrated violence from being prosecuted, even to the extent that the current mens rea regime would otherwise allow. *See e.g.* JOANA SCHWARTZ, SHIELDED: HOW THE POLICE BECAME UNTOUCHABLE (2023); Human Rights Watch, *Shielded From Justice: Police Brutality and Accountability in the United States* (1998).

[13] Though, again, these are additional barriers which prevent survivor-defendants from receiving further mitigation. *See, e.g.,* Tracey Renee McCarter & Samah Sisay, *Prosecutors Must Use Their Immense Discretion to End the Criminalization of Survivors of Gender-Based Violence Who Act in Self-Defense,* 26 CUNY LAW REV. 206 (2023); Liz Komar & Alexandra Bailey, *Sentencing Reform for Criminalized Survivors: Learning from New York's Domestic Violence Survivors Justice Act,* THE SENTENCING PROJECT (2023).

[14] MODEL PENAL CODE § 2.02(2) (AM. L. INST. 1962).

generally.[15] While this Article will focus on criminal law, the distinction between harms inflicted intentionally and unintentionally is central to everything from the international law of war,[16] to corporate law,[17] to anti-discrimination law,[18] to the private law of tort.[19]

This Article questions the soundness of that assumption. It argues that it is possible, as in many cases of reckless or negligent police-perpetrated homicide, that the callous agent who causes some harm recklessly, or even negligently, while indifferent to the possible harmful consequences of their actions, ought to be subject to *more liability* than many cases of run of the mill purposeful wrongdoing, where agents frequently perform the same wrongful act intentionally but reluctantly. In fact, I will argue that the traditional justifications advanced in defense of the assumption that intentional harms should be subject to more liability than unintentional harms – from both the perspective of retributivist theorists concerned with desert and deterrence theorists concerned with dangerousness – actually entail this opposite result.

The claim that we should reject one of the foundational assumptions of modern criminal law's mens rea regime may sound radical. But the basic motivating idea is actually quite intuitive. This Article makes the case that there is an important class of purposeful wrongdoer – the reluctant wrongdoer who engages in intentional wrongdoing as a necessary means to some further goal, but who is also committed to attending to and pursuing alternative means toward that goal, even when those alternative means are costly – who should be subject to less criminal liability than particularly callous wrongdoers who commit the same crimes recklessly or negligently, but who lack such avoidance commitments. Even when the reluctant defendant's reason for acting are insufficient to justify their conduct, and so not the kind of reluctance that would merit a necessity or lesser evils defense, the fact that a defendant committed a wrongful act only as a last resort ought to be a central feature in our calculations of how culpable, and how dangerous, a particular defendant is, and thus play a central role in the construction our mens rea regime.[20]

The article proceeds in five parts. In part I, I unpack the underlying justification for higher criminal liability for intentional wrongdoing, in order to better evaluate whether this justification is strong enough to support the intuitive mismatch

---

[15] *See generally* Paul H. Robinson & Markus D. Dubber, *The American Model Penal Code: A Brief Overview*, 10 NEW CRIM. L. REV. 319, 319-20, 326-29 (2007) (showing that over two-thirds of states have adopted the Model Penal Code (MPC) in whole or in part, and that "even within the minority of states without a modern code, the Model Penal Code has great influence").

[16] *See, e.g.,* Thomas Nagel, *War and Massacre*, 1 PHIL. & PUB. AFF. 123 (1972).

[17] *See, e.g.,* DGLC 102(b)(7) (providing a "raincoat" provision which shields corporate officers and directors' personal liability for negligent, but not intentional, wrongdoing).

[18] *See, e.g., Washington v. Davis,* 426 US 229 (1976).

[19] *See*, *e.g.,* ARTHUR RIPSTEIN, PRIVATE WRONGS 39-51 (2016).

[20] This Article will focus primarily on mens rea requirements involved in criminalization and will not address in detail related mental state issues that arise in sentencing and parallel proposals for sentencing reform. Sentencing takes place in the shadow of the law. While avoidance commitments should also play a role in sentencing guidance, such guidance will be of at best limited use absent reform to the mens rea standards which set the sentencing bands within which such guidance is to be applied. *See infra* notes 24, 25, 26, and accompanying text.

in liability between especially callous reckless or negligent homicides on the one hand and reluctant purposeful homicides on the other.

I argue that the most plausible and popular justification is grounded in what is often called the *doctrine of double effect*, demonstrated most famously by the philosophical 'trolley problem.' Many people intuitively feel an important difference between the agent who knowingly causes the death of a bystander by switching a trolley's tracks in order to save the lives of five others who would otherwise have been struck and killed by a runaway trolley car, and the agent who, for the same reason, purposefully causes the death of a bystander by pushing them onto the tracks to stop the trolley car. One explanation for this difference in intuition is that the agent who purposefully pushes the bystander onto the tracks is committed to their victim's death in a way the knowing, reckless, or negligent agent is not. In the first case, if the bystander escapes, the knowing or reckless agent will be relieved. In the second, the purposeful agent will have to drag them back to the tracks. This commitment to 'track the harm' to the victim across various counterfactual circumstances demonstrates both a more culpable ill-will toward the victim as well as a more dangerous set of dispositions. Callous agents who tolerate harm to victims without intending it can be deterred by simply 'building a fence' around the victim, whereas agents who make harm to a victim their 'conscious object' will not be so easily deterred. Despite new barriers, they will still continue to pursue the harm to their victim, even if they do so reluctantly, so long as that harm is a necessary means to their ultimate goal.

In part II, I push back against this traditional justification and defend a novel account of culpability, centered around reluctance, rather than purpose. I argue that the intentional commitments and dispositions of the reluctant purposeful agent actually reflect more concern for others, not less, than the commitments and dispositions of the callous agent. Although it may be true that the callous agent is committed to better dispositions in counterfactual cases where we place barriers between the victim and the agent, the reluctant purposeful agent is committed to more exculpatory dispositions to avoid wrongdoing in another important but overlooked set of cases where new alternative means become available. In the case of Browne and Williams' study, for example, a key reason survivor-defendants provided to researchers for having committed homicide was that homicide appeared to be a necessary means to the defendants' ends. Absent policy measures like emergency shelters or legal avenues for unilateral divorce, the death of their spouse was perceived as the only means available to them to escape the 'entrapping life situation' in which they found themselves. Had such alternative measures been available, the reluctant purposeful wrongdoer would not have committed the criminal wrong. Whereas the callous negligent or reckless agent like Chauvin, who does not think that his victims' lives matter, lacks such avoidance commitments. They will not be disposed to search for or adopt more inconvenient alternative means, like de-escalatory police tactics, that do not have the consequence of harming others.

Attending to this richer set of dispositional commitments reveals that the standard defense of treating intentions as more liable does not automatically follow from looking at an actor's intentional commitments to harm or help others in various counterfactual circumstances. Instead – whether implicitly or explicitly – the standard

mens rea hierarchy reflects a prioritization on the part of the law toward building fences to deter callous wrongdoers over the creation of new legal or policy alternatives for reluctant purposeful wrongdoers.[21]

In part III, I use recent advances in social psychology, particularly discussions of institutional design in the context of policing, to argue that, as a matter of empirical fact, this prioritization of fence-building over barrier-breaking is a mistake. It is often much easier to create new policy or legal alternatives for reluctant wrongdoers who are cornered into criminal activity through violence, poverty or other environmental contexts as their only means to material security than it is to protect innocent people from callous agents who will tolerate a risk of harm to others as a consequence of pursuing the easiest path to their goals, even when other slightly more costly alternative means with better consequences for others become available.

Marshalling this same empirical evidence, I also make the case that instances of reluctant purposeful wrongdoing are not the exception, but rather commonplace within the class of purposeful wrongdoing. While the case of survivor-defendants and police homicides make stark the problems with the current MPC mens rea regime, and are the site of some of the gravest injustice resulting from that regime, the problem of mismatch is likely widespread across the criminal justice system more generally, with purposeful defendants more likely to exhibit avoidance commitments and reckless/negligent agents less likely, across a wide range of criminalized activity. The assumption that intentional wrongdoing merits more criminal liability can – and has – been used to justify harsher treatment of a wide variety of crimes committed purposefully by defendants who were driven to commit those crimes out of poverty, abuse, and other forms of social marginalization. Whereas it has been used to avoid or reduce criminal liability by reckless or negligent defendants in positions of social power, such as white-collar defendants or landlords who cut corners on safety regulations to cut down on costs, who commit non-intentional crimes of convenience while unwilling or unmotivated to take easily available options to avoid harming their victims.

Parts IV and V show how the Article's novel dispositional analysis of culpability has important doctrinal upshots, by detailing the ways that attending to 'avoidance commitments' can help us amend existing homicide law (and by extension, criminal law more generally) to better apportion liability for reluctant purposeful agents on the one hand, and callous negligent and reckless agents on the other.

The crux of the problem, I argue in Part IV, is that the current doctrinal regime for mitigation and aggravation in criminal law, while intended to augment and fine-tune the PKRN hierarchy, still encodes an implicit assumption about the special blameworthiness of intentional wrongdoing. The premeditation requirement for first degree murder, mitigating doctrines like extreme emotional disturbance and provocation-passion, even relatively more recent affirmative defenses like battered persons syndrome and coercive control and legislative sentencing reforms for

[21] A strategy often referred to as 'target hardening' in the criminology literature. See *infra* note 81 and accompanying text.

domestic violence survivors, all tend to understand mitigating and excusing circumstances as those where the agent is not fully responsible for their actions. These "diminished responsibility" excuses are designed to capture cases where the defendant's action does not reflect their deliberate choices, by picking out cases where the defendant's rational will is overwhelmed by an emotion, so that they are alienated from the resulting action, and so not 'really' or 'fully' intending. But such excuses are orthogonal to, and so often fail to include, cases of reluctant agents whose actions *are* fully deliberate and so fully intentional, but who are committed to engaging in wrongdoing only as a last resort. This mismatch between the source of their reduced culpability, and the available legal models for full or partial excuse, forces reluctant purposeful defendants to choose between either denying their agency through defenses such as battered person syndrome (a defense which, as critics have noted, often problematically frames survivors' experiences as pathological) or else face maximal criminal liability for first degree murder.

After diagnosing the shortcomings of current doctrine, Part V shows how the Article's analysis of the respective culpability and dangerousness of the reluctant purposeful actor and the callous reckless or negligent actor in terms of avoidance commitments can be harnessed to develop an alternative mens rea doctrinal regime. It provides a framework for how legislatures could craft a counter-factual test for avoidance commitments that could be used to augment, or even replace, the familiar four PKRN states of the American mens rea regime. I show how such a system would better track defendants' underlying culpability than our current system, without forcing the fact-finder to make problematic discretionary normative judgments about the quality of a defendant's motives (a common difficulty faced by recent scholarly proposals for mens rea reform). And I show how such a regime could be formulated in terms of counter-factual concepts which we have every reason to think jurors would be as good or better at assessing than the mental states of the current PKRN regime.

Before moving on, a word about the choice of cases. Much of the Article's discussion will center around the two illustrative cases of police-perpetrated reckless and negligent homicides on the one hand, and reluctant purposeful homicides committed by defendant survivors of family abuse, domestic abuse, or trafficking, on the other. I will suggest that a central feature of both these cases – for understanding both the seriousness of Chauvin's offense, and the need for mitigation in cases of survivor-defendants, is the presence or absence of avoidance commitments. In the case of survivor-defendants, a crucial feature of many cases is that the defendant engaged in homicide only as a last resort. The fact that so many survivor-defendants go to such great lengths to avoid homicide – that survivor-defendants try going to the police, social workers, family members, or shelters, sometimes even attempting suicide, before resorting to killing their partner, manifests a deep commitment to avoiding the taking of another human life which the law should take into account.[22] In contrast, a central part of what makes Chauvin's case so morally abhorrent, and so dangerous, is that there were so many other options available for achieving his goal of

[22] This feature is particularly stark in the case of State v. Norman, 378 S.E.2d 8 (N.C. 1989), the most prominent example featured in most criminal law casebooks.

arresting George Floyd that would not have risked the same harm to Floyd, and which Chauvin could easily have availed himself of, but choose not to.

One might worry that the exceptional nature of these cases makes them imperfectly suited to a critical discussion of the importance of purpose in criminal law generally. Both cases introduce complications beyond the agent's place in the PKRN hierarchy, by implicating a variety of other potential factors that might be relevant to our judgments about the relative liability appropriate to each defendant. Purposeful homicides committed by survivor defendants, for instance, often implicate issues of immanence and self-defense, provocation, and diminished responsibility due to extreme emotional disturbance or the psychological effects of trauma. Non-purposeful homicides committed by police officers like Chauvin frequently implicate issues of abuse of power, racial discrimination, and civil rights violations.

The choice to focus on these cases is in part *because of* these complications, not in spite of them. Though exceptional in certain respects, homicide law is also of particular importance for the criminal law theorist because, due to the severity of the crime, it is the place where mens rea doctrine is developed in the most sophisticated manner. In particular, more than for most other categories of offences in criminal law, criminal homicide regimes are not blind to the fact that the rigid ordinal ranking of intentional harms as more liable than harms of recklessness or negligence may fail to track the underlying culpability and dangerousness of some offenders. By focusing on homicide, and on the sorts of complicated real-life cases where avoidance commitments are often caught up with issues of power dynamics, self-defense, provocation, or diminished responsibility, these examples allow for a discussion of why even these more nuanced doctrinal resources for mitigating and aggravating mens rea carveouts in criminal law, along with existing affirmative defenses such as self-defense, duress, or necessity, are still insufficient, even when properly applied, to allow for appropriate mitigation for many reluctant purposeful defendants, and to help see more clearly the shape that a more successful doctrine might take.

Still, as I will detail over the course of the Article, the existence and importance of avoidance commitments generalizes beyond survivor-defendants to many other kinds of purposeful defendants and many other crimes. Perhaps most importantly, avoidance commitments will often be manifested by purposeful defendants in a wide variety of economic crimes, who engage in such crime as a last resort because of a lack of other viable non-criminal social alternatives. While economic crimes like fraud or larceny are not graded by culpability, the hierarchy still shapes the way they are criminalized, with pernicious effect. Like larceny and fraud, most economic crimes have a minimum mens rea requirements of knowledge or purpose. If I am right about the importance of reluctance, and the number of purposeful defendants who manifest such reluctance, relative to reckless and negligent defendants, such mens rea minimums will mean that many of the most culpable and dangerous offenders engaged in reckless or negligent wrongdoing will not be subject to *any* criminal liability on the current mens rea regime. Whereas many of the least culpable defendants who engage

in reluctant purposeful economic crimes as a last resort will always be subject to criminal liability.[23]

Finally, while the focus of this Article is on criminalization, it is worth noting that statutory choices about the mens rea hierarchy are frequently enormously influential, when not outright outcome determinative, when it comes to sentencing as well.[24] For crimes with a minimum mens rea of willfulness, for example, callous reckless agents will not be sentenced at all. Where crimes are graded by mens rea, the mandatory minimum of one grade (such as first degree homicide for deliberate and premeditated purposeful killings), is often higher than or equal to the mandatory maximum of the lower grades (such as second- or third- degree homicide for depraved heat recklessness).[25] Even where grading does not establish a strict lexical ordering of sentencing, the mens rea hierarchy produces higher sentencing bands, as well as higher presumptive sentencing baselines in sentencing guidance,[26] such that sentencing frequently replicates the same mens rea ordering which one finds in criminal statutes.

I. JUSTIFICATIONS FOR THE TRADITIONAL MENS REA HIERARCHY

Given the ways the current legal regime appears to "mis-sort" the liability of reckless and negligent police homicides relative to reluctant purposeful homicides like those surveyed by Browne and Williams, it is worth considering the justifications for the status quo.[27]

---

[23] *See, e.g.,* MODEL PENAL CODE § 2.02(5) (AM. L. INST. 1962) (providing culpability substitution principles so that 'more culpable' states, like purpose, can always substitute for 'less culpable' states like negligence or recklessness, though not vice versa).

[24] A detailed discussion of the ways in which assumptions about the special blameworthiness of purpose shapes sentencing is outside the scope of this Article. However, many of the points made here about the need for more attention on reluctance in the context of criminalization apply with as much or greater force to sentencing practices. My own view is that reluctance needs to play a central role in our mitigation practices at *both* the adjudicatory and sentencing stages, and I hope to explore the parallel problems of sentencing, and the prospects for sentencing reform, more fully in future scholarship. I discuss some of the reasons why leaving the mitigating role of reluctance entirely to sentencing is insufficient in more detail *infra* in Part V. Intentional Commitments, Motives, and Mens Rea Reform.

[25] In Minnesota, for example, where Chauvin was charged, depraved heart reckless homicide, categorized as murder in the third degree, carries a maximum sentence of imprisonment for not more than 25 years, whereas first degree murder carries a mandatory minimum of life imprisonment. *See supra* notes 5 and 6. In the Federal code, the *minimum* sentence of life imprisonment for first degree homicide is the same as the *maximum* upper bound for second degree depraved heart reckless homicide. *See supra* note 8.

[26] In the federal sentencing guidelines, for example, the base offense level of first-degree homicide is 43, with a corresponding recommended sentence of life. The base offense level for second degree murder is 38 with a presumptive sentencing range of 235-293 months. The base offense level for reckless involuntary manslaughter is 18 with a presumptive sentencing range of 27-33 months, and negligent involuntary manslaughter is 12 with a presumptive sentencing range of 10-16 months. U.S. SENT'G GUIDELINES MANUAL §2A1.1 (U.S. Sent'g Comm'n 2023). For the offense of Felony Murder, where mens rea does not determine a presumptive sentencing range, the guidelines still suggest that "[i]f the defendant did not cause the death intentionally or knowingly, a downward departure may be warranted." U.S. SENT'G GUIDELINES MANUAL § 2A1.1 cmt. n.2(b) (U.S. SENT'G COMM'N 2023)).

[27] Recent empirical work also suggests the standard PKRN mens rea hierarchy is less intuitive to the average lay juror than is traditionally assumed. *See* Francis X. Shen, Morris B. Hoffman, Owen D. Jones

There are, broadly speaking, two kinds of defenses on offer: one broadly retributivist in nature, that intentional actors are, at least *ceteris paribus*, more culpable than non-intentional actors for the same harm; the second broadly consequentialist in nature, concerning the degree of dangerousness of intentional, as opposed to unintentional, wrongdoing. Both of these arguments typically rely on some version of what has been historically known as *the doctrine of double effect*.

The doctrine of double effect has a long history, dating back to St. Thomas Aquinas' discussion of self-defense in the *Summa Theologica*.[28] According to the doctrine, when an action has two effects (such as preventing harm to yourself, and causing harm to another), it matters morally whether the actor intends only the first effect, or both.

In contemporary discussions, the doctrine is most famously and vividly illustrated by the "Trolley cases" developed by Phillipa Foot and Judy Thomson.[29] Many people believe there is nothing morally problematic about switching the tracks of a runaway trolley to save five people who would otherwise be run over, even if, in doing so, you know that you will cause the death of an innocent bystander on the second pair of tracks. In contrast, many people believe there is something deeply problematic about pushing that same innocent bystander onto the tracks in order to save the same five people.[30]

It can be puzzling to explain our difference in intuition in such cases, given that the agents' actions will both cause equal harms, and the agents are both equally aware of those harmful consequences. The moral difference between the two cases, proponents of the doctrine of double effect argue, is that in the one case, the harmful effects of the agent's actions are intended or purposeful, whereas in the other case, the harmful effects are merely foreseen but unintended consequences. The intending of the effect seems to be involved in explaining the difference in moral import between these two cases.

There is ongoing debate, among ethicists, about how, and whether, the intentions of an agent could make a difference as to the moral permissibility of an otherwise identical action with identical effects. However, while there is active disagreement about whether the doctrine of double effect should be accepted with respect to the *moral permissibility of actions*, there is far more widespread acceptance of the view that the doctrine of double effect is importantly on the right track when it

---

& Joshua D. Greene, *Sorting Guilty Minds*, 86 N.Y.U. L. REV. 1306, 1339-43 (2011). While the study authors have suggested that this may be the result of failures of jury comprehension, I suggest elsewhere that another possibility is that jurors reject the underlying normative arguments for the hierarchy. *See* Gregory Antill, Note, *Fitting the Model Penal Code Into a Reasons-Responsiveness Conception of Subjective Culpability*, 131 YALE L. J. 1346 (2022). This article is in many ways an expansion and continuation of the arguments begun there.

[28] THOMAS AQUINAS, SUMMA THEOLOGICA (II-II, Qu. 64, Art.7).

[29] Phillipa Foot, *The Problem of Abortion and The Doctrine of Double Effect in* VIRTUES AND VICES (1978); Judith Jarvis Thompson, *The Trolley Problem*, 94 YALE L.J. 1395 (1985).

[30] Thompson, *supra* note 29.

comes to the *evaluation of the blameworthiness of agents*.[31] Indeed, among those who deny the role of an actor's intentions in determining the moral permissibility of an action, the most common explanation for the mistake on the part of their opponents is that they are conflating questions of the moral blameworthiness of the actor with the permissibility of the action.[32]

Still, among advocates of the so-called *agent evaluative-version* of the doctrine of double effect, there is debate about the underlying source of the doctrine of double effect, with some arguing that the intention or absence of intention has a *direct* effect on the actor's blameworthiness, where other theorists argue that the effect on blameworthiness is *indirect*.[33] This split between direct and indirect accounts maps onto two different strategies for defending the PKRN mens rea regime.

I.A. Insufficient Concern as a Basis for the PKRN Mens Rea Hierarchy

The most common *indirect-blameworthiness* account of intention comes from a family of views often labeled 'Insufficient Concern' or 'Quality of Will' accounts of culpability. On this picture, the basis of assessment for an agent's culpability with respect to some wrongful action is not the agent's proximate PKRN mental states (whether the action was intentional, as opposed to merely knowing, reckless, or negligent) but rather the more distal reasons for acting, which their actions, performed purposefully, knowingly, recklessly, or negligently, evince. The weight the agent gives these reasons helps constitute the agent's *quality of will* – the agent's "take on the world and what is important or worthwhile or valuable in it."[34]

There are certain features of the world which morality – or the law – requires agents to be concerned about when deciding how to act. In particular, morality and the law typically demand that agents show concern for other people. On this picture, agents are blameworthy to the extent to which their concern deviates from the concern of a lawful agent.[35] Agents whose behavior manifests a high degree of concern for the legally protected interests of other people evince a good quality of will, whereas agents whose behavior manifests a lesser degree of concern for the legally protected interests of other people evince a more culpable quality of will, to the degree that their concern falls short of the ideal agent.[36]

---

[31] *See e.g.* Alexander Sarch, *Double Effect and The Criminal Law*, 11 CRIM. L. & PHIL. 453 (2017).

[32] *See* T.M. SCANLON, MORAL DIMENSIONS: PERMISSIBILITY, MEANING, BLAME 13 (2008) ("the difference between causing harm intentionally and doing so negligently . . . . is not a difference in *permissibility*. Both are generally impermissible. The difference between them lies, rather, in the kind of fault that is involved when an agent acts impermissibly in these ways."). Defenders of the culpability version of the doctrine of double effect and its role in criminal law include Sarch, *supra* note 31, MICHAEL MOORE, CAUSATION AND RESPONSIBILITY 48 (2009), and Dana Kay Nelkin and Samuel C. Rickless, *The Relevance of Intention to Criminal Law,* 10 CRIM. L. & PHIL. 745 (2016).

[33] This framing is from Nelkin & Rickless, *supra* note 32.

[34] Pamela Hieronymi, *Responsibility for Believing,* 161 SYNTHESE 357, 361-362 (2008).

[35] *See e.g.* LARRY ALEXANDER AND KIMBERLY KESSLER FERZAN, CRIME AND CULPABILITY (2009); GIDEON YAFFE, AGE OF CULPABILITY (2018).

[36] *See* ALEXANDER & FERZAN *supra* note 35.

As the philosopher P.F. Strawson, the progenitor of the contemporary quality of will account puts the point, reactive attitudes such as blame are "essentially reactions to the quality of others' wills towards us, as manifested in their behaviour: to their good or ill will or indifference or lack of concern."[37] Such assessments "rest on, and reflect, an expectation of, and demand for, the manifestation of a certain degree of goodwill or regard on the part of other human beings towards ourselves; or at least . . . an absence of the manifestation of active ill will or indifferent disregard."[38] When an agent commits some crime by causing harm to another person, they are culpable to the extent that this action manifests a lack of sufficient concern or goodwill toward the person being caused harm, as the law or morality demands.

In focusing on the agent's reasons for actions, the quality of will account is concerned with more fine-grained features of the agent's subjective psychology than the PKRN mens rea categories. Just as two agents who perform the same action might do so with different mens rea on the PKRN regime (such as purpose or recklessness) and so be differently culpable, on the quality of will account, two agents might perform the same action purposefully but be differently culpable based on their differing reasons for acting purposefully.

Consider, for example, an agent who causes me injury by shoving me. As Strawson suggests, the blame we assign to the agent tracks not, in the first instance, psychological facts about whether the action was intentional or unintentional, but rather more distal psychological facts concerning the degree of concern which their intentional actions evince. As Strawson says, though in each case the "pain may be no less acute," it matters if they shove me "while trying to help me," say by pushing me out of the way of incoming traffic, if they are shoving me simply to get me out of their path "in contemptuous disregard of my existence," or if they are shoving me out of a "malevolent wish to injure me."[39]

Still, while it is an agent's concern – or lack of concern – for others which is, strictly speaking, the proper object of assessment for assigning culpability on the quality of will account, an agent's intentions is indirectly relevant to culpability because it provides important *evidence about* the agent's underlying concern or lack of concern. As T. M. Scanlon explains:

> [an agent's intention] . . . tells us something about an agent's view of the reasons bearing on his or her action…what it is that she is doing intentionally tells us what she believes about her situation and the likely effects of her action . . . . it also tells us something about how she evaluates these factors – which she sees as reasons for acting the way she plans to act, which as costs to be avoided if possible, which as costs to be borne, which as inconsequential[40]

---

[37] Peter Strawson, *Freedom and Resentment*, 48 PROC. BRIT. ACAD. 1 (1962), *reprinted in* PERSPECTIVES ON MORAL RESPONSIBILITY 45, 48-50 (John M. Fisher & Mark Ravizza eds., 1993).
[38] *Id.* at 48-49.
[39] *Id.*
[40] SCANLON, *supra* note 32.

For proponents of the quality of will account, then, facts about whether an agent did or did not intend some action (say, causing the death of the other person) are not direct grounds for culpability, but rather *proxies* for underlying failures of reasoning or lack of concern on the part of the agent. If, for example, an agent intentionally or knowingly or recklessly speeds to get home more quickly, they must have judged that getting home was more important than the risk of harm they placed upon other drivers, and so demonstrated a culpable lack of concern, or absence of good-will.

One way to justify the PKRN hierarchy, then, would be to show that the PKRN ordering of culpability proxies tracks the relative underlying quality of will of the actor.[41] That is, to show that, for a given harm, the agent who causes that harm purposefully or intentionally manifests a worse quality of will than the agent who causes that same harm knowingly, who in turn manifests a worse quality of will than the agent who acts recklessly or negligently when a risk of the harm is foreseen (or foreseeable).[42]

Many criminal law theorists have argued for precisely such a claim. The crux of the argument involves the observation that the intentional agent, unlike the knowing, reckless, or negligent agent, must make wrongdoing their "conscious aim."[43] By "aiming at evil" the purposeful agent will necessarily evince more ill will than the agent who is merely "tolerating" such evil.[44] As Kimberly Ferzan has explained, this "aiming at evil" response can be used to justify the PKRN scheme:

> Purpose is more culpable than knowledge because purpose entails aiming at the wrong, while knowledge entails toleration of the known wrong . . . . As for recklessness, the actor's epistemic uncertainty means that the actor does not identify with or choose the wrong, in the same manner as knowing and purposeful actors. Hence, culpability establishes the defendant's willingness to identify herself with wrongdoing. The more the wrong is part of the actor's reasons for acting, the more culpable she is and the more she deserves to be punished[45]

---

[41] *See* Antill, *supra* note 27, for further argument that the MPC Mens Rea states function as proxies for underlying quality of will.

[42] This view is widespread. *See, e.g.*, Kenneth W. Simons, *Rethinking Mental States,* 72 B.U. L. REV. 463, 490 (1992) ("The reigning hierarchy often works fairly well in translating underlying normative approaches [to] blameworthiness . . . into doctrinal requirements."); Douglas Husak, *"Broad" Culpability and the Retributivist Dream*, 9 OHIO ST. J. CRIM. L. 449, 454-55 (2012) ("Ceteris paribus, a defendant who performs the actus reus of a crime purposely is more blameworthy than one who acts knowingly, who in turn is more blameworthy than one who acts recklessly, who in turn is more blameworthy than one who acts negligently, who in turn is more blameworthy than one who is strictly liable because he acts with no culpability at all."); Gideon Yaffe, *The Point of Mens Rea: The Case of Willful Ignorance*, 12 CRIM. L. & PHIL. 19, 19 (2018) (justifying the place of willful ignorance, and the PKRN mens rea hierarchy more generally, through its function as proxies for the agent's underlying reasons-responsiveness).

[43] MODEL PENAL CODE § 2.02(2)(a)(i) (AM. L. INST. 1962).

[44] THOMAS NAGEL, THE VIEW FROM NOWHERE 181 (1986).

[45] Kimberly Kessler Ferzan, *Don't Abandon the Model Penal Code Yet! Thinking Through Simons's Rethinking*, 6 BUFF. CRIM. L. REV. 185 (2002).

As Ferzan illustrates, the PKRN hierarchy may appear to map neatly on to Strawson's hierarchy of ill-willed actors. Morality and the law insist that we show a high degree of concern for the well-being of others, taking the fact that our actions risk harm to other people as a strong (and typically decisive) reason to refrain from acting.[46] The reckless agent who willingly risks harm to another must treat the fact that their action risks harm as a weaker reason, demonstrating culpable "disregard" for their well-being. The Knowing agent demonstrates even more "indifferent disregard" for the other's well-being. When the reckless agent treats the fact that another person would be harmed as a reason, with a certain weight, to refrain from acting, that weight will be discounted when deciding how to act by the subjective probability they assign to the risk of harm. The agent who knows that they will harm the victim has no such discounting, and so must have valued their victim's well-being even less.[47]

Both the reckless agent and knowing agent, then, demonstrate an *insufficiently high degree of concern* for the victim's well-being. They grant their well-being some positive normative weight in their reasoning, but less than they should. Still, the argument goes, if we measure lack of concern by the difference between this positive weight and the ideal weight, the purposeful agent will always demonstrate an even greater ill-will, or lack of regard, for the well-being of others. Since the purposeful agent is aiming at harming the other, they are not simply undervaluing other peoples' well-being, but rather treating the harm as a positive reason for acting (and so assigning a *negative value* to their victim's wellbeing).

Thus the adoption of a *quality of will* account of culpability can provide us with a normative justification, in terms of culpability, for treating purposeful agents who act out of a "malevolent wish to injure" as worse than callous agents (whether knowing, reckless, or negligent) who treats another with "contemptuous disregard."[48] The fact that an action, A, will cause another harm ought to be treated as a reason, with some negative weight, R, against acting. A callous agent demonstrates, by A-ing, that they 'tolerate' the harm, and that they have assigned to the fact that A will cause harm some smaller normative negative weight less than R. Reckless or Knowing agents who act with "extreme indifference to human life" are the limit case, where they grant zero weight to the other agent's well-being.

In contrast, the purposeful agent, if harming is their aim, treats the fact that the victim will be harmed as a reason with some positive weight, W, in favor of acting. As Thomas Nagel has put the point, the purposeful agent who aims at evil is not just failing to fully appreciate the normative force of the reasons other people give us, but

---

[46] There are, of course, some circumstances where, if the countervailing reasons for acting are sufficiently strong, certain harms can be justified. These are represented in the law through affirmative defenses like necessity and self-defense. *See, e.g.,* MICHAEL MOORE, PLACING BLAME: A THEORY OF THE CRIMINAL LAW (2010).

[47] At least if we hold fixed the force of their countervailing reasons. *See* Antill *supra* note 27; Alexander and Ferzan *supra* note 35.

[48] Scanlon *supra* note 32. In the language of economics, the agent's intentions give us at least some information about their 'revealed preferences' with respect to the foreseen consequences of their chosen action.

rather "swimming head on against the normative current."[49] When the purposeful agent makes producing harm their "conscious object,"[50] or goal, they assign to the fact that they will produce harm a normative weight in a "direction diametrically opposite to that in which the value of that goal points."[51] And to grant a reason some force "opposite" to its actual force is a greater failing than to grant it no force at all. The callous agent is at worst showing an insufficient concern of degree R (the ideal weight, R, less the weight they assigned it, 0). Whereas the purposeful agent is at best showing an insufficient concern of degree R + W (the ideal weight, plus whatever positive weight they have assigned their goal of causing harm.)

I.B. Reluctance, Indifference, and Insufficient Concern

In the previous section, I have described what might be called the "aiming at evil" argument for the PKRN hierarchy. But as I and others have argued in the past, this argument is too quick.[52] The fundamental problem is that agents who intend some action, A, as a means to some further goal, B, need not see the 'evil-making' features of A as a reason to act. In fact, intending to A as a means to B is consistent with seeing A as counting against the overall enterprise and providing a reason (perhaps a powerful reason) not to act. Of course, this *reluctant purposeful agent,* who sees the intended harm as a reason not to act, must see the force of R as outweighed by the perceived value of the final goal, given that they still ultimately form the intention to proceed with the act.[53] So an agent who intends some harm to another as a means to some further goal must at least tolerate the harm, thus revealing a lack of sufficient concern relative to the ideal agent. But the knowing agent who understands that the unintended harm will result as a foreseen side-effect of their actions must also see the force of R as outweighed by the perceived value of their ultimate goal. The perceived disvalue of the harmful effect will thus be outweighed by the perceived value of the agent's goal whether an effect is merely tolerated or intended.

Whether the agent acts knowingly or purposefully is a matter of whether the harm happens to be instrumental, or not, to forwarding their overall plan.[54] But whether an agent treats the harm as a cost or a benefit of their overall plan is independent of whether the harm plays such an instrumental role.[55] Just like the knowing or reckless agent, the purposeful agent may act in spite of the fact that their plan involves harming another, not because of it. It is a matter of moral luck whether the world is structured such that the harm happens to be a means of achieving the

---

[49] NAGEL, *supra* note 44.
[50] MODEL PENAL CODE § 2.02(2)(a) (AM. L. INST. 1962).
[51] NAGEL, *supra* note 44.
[52] *See, e.g.*, Antill *supra* note 27; Kenneth W. Simons, *Punishment and Blame for Culpable Indifference*, 58 INQUIRY: AN INTERDISCIPLINARY JOURNAL OF PHILOSOPHY 143 (2015); ALEXANDER & FERZAN, *supra* note 35; SHELLY KAGAN, THE LIMITS OF MORALITY 128 – 182 (1999); Clare Finkelstein, *The Irrelevance of the Intended to Prima Facie Culpability: Comment on Moore,* 76 B.U. L. REV. 335 (1996).
[53] Putting aside for the sake of argument the possibility of weakness of will. *See* Donald Davidson, *How is Weakness of Will Possible? in* ESSAYS ON ACTIONS AND EVENTS 21 (2001).
[54] *See, e.g.,* Michael Bratman, *Moore on Intention and Volition,* 142 U. PENN. L. REV. 1705, 1706 (1994).
[55] *See, e.g.,* MICHAEL BRATMAN, INTENTION, PLANS AND PRACTICAL REASONS 152-155 (1987); KAGAN, *supra* note 52.

goal, or whether the harm is a mere causal consequence, or side-effect, of pursuing the goal.[56]

Once we see that a purposeful agent who intends some harm as a means to some further goal can both intend to harm another person *and* treat the fact that they will harm the other person as a reason to refrain from acting, we can see that the *quality of will* account of culpability cannot explain why reluctant purposeful homicide, like the survivor-defendant homicide cases described by Browne and Williams, are more culpable than the callous reckless police-perpetrated homicide, like that of Derek Chauvin.[57]

In fact, the *quality of will account* appears to suggest the opposite. According to the quality of will account we have been considering, an agent whose actions manifest an insufficient degree of concern for others is culpable to the degree their lack of concern diverges from the normative ideal. Since the reluctant purposeful agent grants the well-being of the other agent at least *some weight*, where the reckless agent who exhibits "extreme indifference to human life" grants the well-being of the other agent *no* weight, the reluctant purposeful agent will demonstrate *less ill will*, and so be less culpable.

Consider a simplistic Burden vs. Probability of Loss (BPL) model of care, adapted from the law of torts,[58] where the manifest degree of concern will be a function of (a) the strength of the agent's countervailing reasons for acting, B, and (b) the subjective probability, P, they assign to the harm their action may cause.[59] If the practical weight they give to the other person's loss, L, is high enough that, discounted for its probability, PL was greater than the weight they assigned to B, they would not have acted as they did. So, if they *do* act in ways they foresee risk harm to another, to achieve the benefit they are pursuing, the maximum weight, or concern, they can have granted the other person's well-being in their reasoning is B/P.

But as we have seen, many reckless homicides involve a callous agent who is willing to risk grave harm to others for relatively small rewards. This means that even if they grant only a very low probability of harm (let's suppose, a 1% chance), the large loss that they are risking, combined with the negligible benefits for which they are willing to create such a risk (e.g. the extra entry fees the night club owner will receive by locking the fire exit,[60] or the few minutes saved by the arresting officer who chooses to use a chokehold rather than de-escalate[61]) shows that reckless agent cannot have granted much weight, if any, to the well-being of their victim. The fact that they are

---

[56] An instance of moral luck of the kind Tomas Nagel would categorize as 'circumstance luck.' *See* Thomas Nagel, *Moral Luck*, *in* MORTAL QUESTIONS 24 (1979).

[57] I will consider the case of negligence in more detail later on in Part II, *infra* note 91 and the accompanying text.

[58] *See* U.S. v. Carroll Towing, 159 F.2d 169 (2d Cir 1947).

[59] *See* ALEXANDER & FERZAN, *supra* note 35.

[60] Commonwealth v. Welanksy, 55 N.E.2d 902 (Mass. 1944).

[61] *See generally* Tracey Meares et al., *Principles of Procedurally Just Policing*, THE JUSTICE COLLABORATORY AT YALE LAW SCHOOL 41-44 (2018) for a discussion of such tactics. See also the discussion *infra* Part III.

.

willing to risk a life for such a small gain shows that they, in a literal sense, treat life as cheap.

In contrast, if the agent engaging in purposeful homicide has a more subjectively compelling goal, for which homicide was the means, which provides them with a much stronger countervailing reason for acting, their action will be consistent with a much higher degree of concern for the life of their victim. Consider again the case of the survivor-defendant homicide, like the female perpetrated intimate partner homicides where the defendant killed her partner because she believed this to be the only means to escape an "overwhelming and entrapping life situation."[62] Because their goal provides them a stronger reason (a weightier B), their actions are consistent with stronger degree of concern. If, as is plausible, the higher weight of the purposeful agent's B is more than 1/P times stronger than the risk foreseen by the callous reckless agent, the purposeful agent's actions will be consistent with demonstrating a *greater concern* for the victim's well-being in their practical reasoning.

Of course, this is not to say that the reluctant purposeful agent is necessarily blameless.[63] Both the reluctant purposeful agent and the callous reckless agent may fall well short of the degree of concern demanded by either morality or the law. Still, it does seem to show that the quality of will conception of culpability, described above, not only fails to justify a higher criminal liability of purposeful agents over reckless agents, but, in the case of recklessness with extreme indifference to human life and the relatively common case of a reluctant purposeful agent, appears to support treating the callous wrongdoer as *more liable* than the purposeful wrongdoer, not less.

I.C. Reluctance, Indifference, and Intentional Commitments

On the quality of will picture of culpability, culpability is a function of the weight one actually assigns one's various reasons for acting in ways that risk harm to others. As we have seen, it is unable to explain why intentional wrongdoing should be, *ceteris paribus*, more culpable than knowing or reckless wrongdoing. If we treat intentions as indirectly relevant to culpability only insofar as they are relevant to the quality of the agent's actual practical reasoning, there will be no special culpability in purposefully making a harm your "conscious object" to achieve some further goal relative to tolerating that same harm as a collateral consequence to achieve that end. As Alexander and Ferzan put the point, "purpose, too, is a comparison of risk and reasons" and so, at least with respect to analyzing the underlying quality of will, "just a special case of recklessness."[64] And while a purposeful actor may treat the harm they

---

[62] Browne & Williams, *supra* note 9, at 78.

[63] Nor is it to claim that the defendant is partially *justified*. The strength of their positive reasons for acting matters only because of its role in reverse engineering the value the defendant assigns to L. This argument for reduced culpability does not require claiming that the defendant's positive reasons make the act any less wrong, that the victim 'had it coming,' or that the victim's life was any less valuable or worthy of society's protection. *See* Susan Rozelle, *Controlling Passion: Adultery and the Provocation Defense,* 37 RUTGERS L.J. 197, 215 (2005); Joshua Dressler, *Provocation: Partial Justification or Partial Excuse?,* 51 MOD. L. REV. 467 (1988); *Compare* Mitchell Berman and Ian Farrell*, Provocation Manslaughter as Partial Justification and Partial Excuse,* 52 WM. & MARY L. REV. 1027 (2011).

[64] ALEXANDER AND FERZAN, *supra* note 35.

cause as a reason to act, they might also make harm their intentional aim in spite of treating it as a reason not to act, outweighed by the force of their countervailing reasons to achieve their ultimate goal.

Crucial to this argument from quality of will, however, is that it adopts a picture of an agent's culpability where culpability is a function of the agent's *operative reasons*: the reasons an agent was actually weighing in their practical reasoning when choosing how to act.[65] In contrast, a growing number of philosophers and criminal law theorists have argued instead that the relevant basis of culpability should be a broader assessment of an agent's reason-responsiveness. Agents are culpable not (only) based on their actual lack of concern for the relevant reasons while acting, but also based on what those actions reveal about what Gideon Yaffe has described as their more general 'modes of reasoning.' These more general 'modes of reasoning' include:

> modes of recognition, weighing, and response to reasons [which] consist in dispositions of the agent to display patterns in deliberations, or exercises of practical reasoning…. the distinctive feature of deliberation is the way in which those factual beliefs lead to conceptions of prospective acts as supported or unsupported by reasons, and, in turn, to choices to act. It is in those guiding conceptions of facts as reason-giving that we find the agent's modes of recognition and weighing of reasons.[66]

And whereas intentions are at best indirect and defeasible evidence about what the agent takes as a reason to act and those reasons' force, the fact that an agent intends an act – rather than foresees it – plausibly has a more direct bearing on the broader "dispositions of the agent to display patterns in deliberation or exercises of practical reasoning."[67]

As Michael Bratman has shown, intentions cannot be reduced to an agent's reasons, beliefs, or desires.[68] Part of what it is to have an intention is to have certain *normative commitments* to bringing about the intended state of affairs, above and beyond the normative commitments one has in virtue of the weight one gives various reasons for acting.[69]

This is perhaps easiest to see in the class of cases action theorists describe as 'Buridan Ass cases' named after the proverbial donkey torn between two equally appetizing bales of hay.[70] Suppose that a similarly situated human agent is not so

---

[65] Thanks to Pamela Hieronymi for helping me to see the importance of this point.

[66] YAFFE, *supra* note 35. *See also* Sarch, *supra* note 31 at 464 (2017) ("The main difference between (i) intending bad or wrongful states of affairs—which for simplicity I henceforth refer to as intended harms—and (ii) merely foreseen ones is a matter of one's commitment to them.").

[67] *Id.*

[68] *See* BRATMAN, *supra* note 55.

[69] *See generally* G.E.M. Anscombe, INTENTION (1957); BRATMAN *supra* note 55; Sarch, *supra* note 31; Gideon Yaffe, *Criminal Attempts*, 124 Yale L. J., 1 (2014);); Gregory Antill, *Epistemic Freedom Revisited*, 197 Synthese 793 (2020).

[70] *See, e.g.,* MICHAEL BRATMAN, FACES OF INTENTION 209-244 (1999).

asinine as to be frozen indefinitely between two such equally appealing options, but instead eventually acts and moves in the direction of one over the other. If an agent has equal reasons with equal weight to A and to B, her reasons and the weight of those reasons cannot explain why she As, rather than Bs. Intentions are the kind of mental states that can help do the explanatory work. They do so by providing the agent with extra normative pressure to bring about the intended state of affairs, above and beyond the weight provided by the strength they grant the prior reasons for those states of affairs. Though the agent has equal reason to A and to B, their intention to A commits them to A-ing in a way they are not committed to B-ing.[71]

In particular, the agent who intends to bring about some state of affairs, either as a means or an end, will be committed to *promoting* the state of affairs in a way that an agent who simply foresees the state of affairs will not.[72] Agents who intend to A are committed to acting in ways that are consistent with A – they will be committed to refraining from intending other projects that are incompatible with A.[73] Agents who intend A are also committed to *tracking* A.[74] If circumstances change, and it looks like A will not occur, an agent who intends to A is rationally committed to adapting new sub-plans to ensure that A obtains.

To see how these tracking dispositions might affect assessments of the agent's reason-responsiveness, we can return to the trolley case from the beginning of the section. Consider again the difference between the agent who knowingly causes the death of a bystander by switching tracks in order to save the lives of five others who would otherwise have been struck and killed, and the agent who, for the same reason, purposefully causes the death of a bystander by pushing them onto the tracks.

If we look only at the practical reasoning of the two agents, both agents look very similar. Both agents are weighing the value of the five lives saved against the one life lost. And both agents' choices are consistent with giving the same weight to the value of both the five and the one and concluding that the lives saved are worth the cost. To borrow a distinction from Michael Bratman, there is no difference in 'what is chosen' by the two agents, even though there are differences in what is intended.[75] In this sense, agents will each grant the life of the victim the same weight in their reasoning, and so demonstrate the same 'quality of will' toward the victim. A quality of will account, then, cannot explain our intuitions about the moral difference between the two cases.

If we look at the differing *intentional commitments* of the two agents, however, and the consequences for their various dispositions to patterns of practical reasoning

---

[71] *Id.*

[72] *See* BRATMAN, *supra* note 55. The importance of these commitments to criminal law is particularly clearly articulated in Sarch, *supra* note 31, at 464, caching out the distinctive culpable commitments of the intentional agent in terms of a commitment to be "motivated to perform variations of his actual conduct, which differ in terms of whether they make [the victim's] death more likely" as well as a "commitment [to take] *further* step to ensure that [the victim] dies."

[73] *Id.*

[74] *Id.*

[75] *Id.*

in varied circumstances, stark contrasts between the degree of concern toward the victim emerge.

If it looks like bystander is escaping in the first scenario, the knowing agent will be relieved. The agent saw the harm to the bystander as a reason not to act all along, and so it will turn out that, luckily, the action resulted in even better consequences than the agent had initially hoped. Since the victim's death was not a means for achieving their goal, the agent has no further reason to act any further to cause the victim's death.

In contrast, if it looks like the victim will escape in the second scenario, the purposeful agent, for whom the victim's death is the means of saving the five, will have to take further steps to prevent the victim from escaping. The fact that the agent intends that the victim die as part of their plan for saving the five means that they must 'track' the first victim's death, revising their intentions and creating subplans to ensure that the death occurs so that their ultimate goal of saving the five is achieved.[76] This is in stark contrast to the driver of the tram who switches tracks and knowingly or recklessly risks the bystander's life.[77] As Philippa Foot notes, in counter-factual cases where the bystander lives, "The [reckless] driver of the tram does not then leap off and brain him with a crowbar."[78]

This "tracking" disposition to drag the fleeing bystander back onto the tracks or "brain him with a crowbar" if he looks like he will survive certainly appears to be a larger failure to respond to the value of human life than a disposition not to drag the bystander back onto the tracks. More generally, a disposition to doggedly pursue another person's death in a wide variety of counter-factual circumstances seems to constitute a distinctive way of failing to appreciate the value of that person's life, or to grant their well-being insufficient normative force in your practical reasoning. If purposeful homicides necessarily manifest a commitment to such counter-factual dispositions, and knowing or reckless homicides do not manifest a commitment to such dispositions, this fact can be constitutive of a larger culpable failure of reasoning on the part of the purposeful agent.

This means that even if the reckless or knowing agent and the purposeful agent engaged in criminal homicide grant the same normative weight to the fact they will cause harm in their practical reasoning, the agent who commits purposeful homicide will have different rational commitments. As Gideon Yaffe has put the point:

> intentions constitute commitments to the conditions they depict by generating special reasons for the intending agent to structure his practical reasoning around those conditions. Because the intending killer's intention depicts another's death, he is under rational pressure to ignore options incompatible with the other's death, and to form

---

[76] *Id.*

[77] *But See* JONATHAN BENNET, MORALITY AND CONSEQUENCES 95-116 (1980) for an argument that these counter-factual differences may be narrower than commonly thought.

[78] Foot, *supra* note 29.

> intentions to take means [which will promote the other's death], among other rational pressures. This tells us something of great significance to the assessment of his criminal responsibility, for it tells us how he employs and directs his distinctive human capacity for self-consciously recognizing and responding to reasons.[79]

Because these commitments, constitutive of intentions, have a role in our dispositions to certain patterns of practical reasoning, they will be of further relevance to a reasons-responsiveness conception of culpability that encompasses such dispositions. As Yaffe argues "the role of intentions in constituting commitments explains . . . why intentions are of such paramount importance to culpability and criminal responsibility" not only because "what an agent intends tells us a great deal about what kinds of considerations he recognizes as giving him reason, and about how he weighs those considerations in his deliberation about what to do" but also because intentions are not just evidence of, but "in part *constitutive of* those facts."[80]

The story for why we should treat all purposeful wrongdoers as more liable for a given harm– even the reluctant purposeful wrongdoer who treats the harm as a reason against acting – is that all purposeful wrongdoers are normatively committed to tracking the wrong across even quite distant counter-factual situations. This commitment is not merely evidence of insufficient concern, it is partly *constitutive of* insufficient concern. It is a commitment or disposition to engage in a pattern of reasoning that constitutes a special kind of insufficient concern for the well-being of others, above and beyond the weight one might give that person's well-being in any given piece of practical reasoning.

This does not change the fact that the callous agent is, in other respects, exhibiting worse reasons-responsiveness than the reluctant purposeful agent. The reluctant purposeful agent still gives a stronger weight to the well-being of their victim than does the callous agent in their practical reasoning, and so demonstrates, in that sense, less ill-will toward the victim than does the callous agent. Nevertheless, consideration of the normative commitments constitutive of intentions can give us a way to see why we might think that the agent who takes the death of another as their "conscious object" and so aims at their harm "even as a means" is necessarily more culpable. One need only hold, as seems at least *prima facie* plausible, that to "structure one's practical reasoning" around the aim of harming another, so that one treats the fact that certain actions would prevent harm as decisive reasons to avoid the action, is necessarily a larger failing than to grant that harm less weight in any particular context. And while this is contestable, it's certainly not a wildly implausible thought.

So far, I have focused on how tracking commitments can be used by a proponent of the PKRN hierarchy to defend the claim that intentional wrongdoing is more culpable, and so deserving of more liability, on retributivist grounds. But it can also be used to defend the PKRN hierarchy on more consequentialist deterrence-based grounds as well. Indeed, the argument is even more straightforward. However

---

[79] Yaffe, *supra* note 35 at 111.
[80] *Id.* at 110. Emphasis added.

plausible it is that tracking dispositions are of primary importance for culpability and desert, it seems clear that at least with respect to dangerousness and deterrence, dispositions matter most. The fact that an agent acts with ill-will, desiring to harm the victim, is relevant to dangerousness primarily because an agent who desires to harm the victim will generally be more disposed to be guided in their actions by that desire. But one can imagine an agent – the 'cautious sadist' – who desired to harm victims, but was never or only very rarely disposed to act on that desire, because the desire is always, across contexts, outweighed by counter-vailing desires to avoid punishment, so that the agent never actually intends any harm. Whatever their culpability, such an agent will be much less dangerous than an agent who desires not to harm, but who exhibits a disposition to engage in patterns of practical reasoning where that motivation not to harm is often outweighed by countervailing dispositions.

And there appears to be a strong argument that the tracking dispositions of a purposeful agent are much more dangerous, and harder to deter, than the dispositions of even a very callous reckless or negligent agent. Because the knowing, reckless, or negligent agent is not aiming to cause harm as a goal, the argument goes, it is relatively straightforward how one protects the vulnerable from harm. When the knowing, reckless, or negligent agent harms another person, say by driving over their property in pursuit of some goal, it is because the path that led them over the property, leading to the harmful collateral consequences, was the most straightforward path to their ultimate end. All one needs to do to deter them is 'create a fence' around the property.[81] Once the action with harmful consequences is more costly than an alternative route, they will go the other way, flowing, like water, toward the path of least resistance.[82]

In contrast, because the purposeful agent has a compelling goal that can only be achieved by harming the victim, their commitment to harming the victim, constitutive of their intention, will dispose them to choose to continue trying to harm the victim, despite obstacles the State may try to put in their path to deter them, until those disincentives are so large as to outweigh the value of the ultimate goal.

As Seana Shiffrin has put the point, "malicious agents are more likely to cause harm than negligent agents who will not double back if harm is avoided."[83] Shiffrin notes that this argument works only against the "determined malicious agent and not the one who suffers a flash of temper that quickly subsides."[84] And this observation perhaps explains part of the standard doctrine for mitigating certain intentional homicides which are not "premeditated" and even further mitigating the subset of those non pre-meditated homicides that are the result of "passion" or "extreme

---

[81] This strategy of deterrence is sometimes referred to as *target-hardening*. For a general discussion of such a strategy, see Meryl & Maurice M. Bell, *Crime Control: Deterrence and Target Hardening*, *in* HANDBOOK ON CRIME AND DELINQUENCY PREVENTION (E Johnson, ed., 1987).

[82] For discussion of this 'economic approach' to criminal liability, see, for example, Gary Becker, *Crime and Punishment: An Economic Approach,* 76 J. OF POL. ECON. 169 (1968); Richard Posner, *An Economic Theory of the Criminal Law*, 85 COLUM. L. REV. 1193 (1985); Keith Hylton, *The Theory of Penalties and the Economics of Criminal Law*, 1 REV. OF L. & ECON. 175 (2005).

[83] Seana Valentine Shiffrin, *The Moral Neglect of Negligence*, *in* 3 OXFORD STUD. IN POL. PHIL. 116 (2018).

[84] *Id.*

emotional disturbance."[85] But this caveat does nothing to limit the apparent force of the tracking argument against cases of reluctant purposeful homicides like the case of female perpetrated intimate partner homicides, since reluctant purposeful homicides by survivor-defendants can still be both "premeditated" and "determined."[86]

Putting together these discussions of both culpability and dangerousness, then, we are now in a position to see how a focus on tracking dispositions thus reveals a way in which the cases of reluctant purposeful homicide described by Browne and Williams might be understood as both more dangerous and more culpable than Derek Chauvin's murder of George Floyd, and similar cases of police-perpetrated reckless or negligent homicide. Had the initial arresting officers been able to secure Floyd in the police car, so that Chauvin's putting Floyd in a chokehold would not have furthered Chauvin's intention to arrest Floyd, Chauvin would not (we can assume) have been disposed to act in ways that would still cause Floyd's death. In contrast, in a case of a female perpetrated homicide where the agent believed that the death of her spouse was a necessary means of escaping her entrapping life situation, she would presumably have continued attempting to commit homicide, even if for some reason the victim escaped harm from her initial actions. Though her motives may be sympathetic, this disposition seems both more culpable and more dangerous. Generalizing from these cases, we can see how a focus on tracking dispositions appears to provide a novel way to defend the PKRN mens rea hierarchy against recent critics on both retributivist and consequentialist-deterrence grounds. Nonetheless, in Part II, I will argue that, despite the apparent strengths of these arguments, they are ultimately unsuccessful. In fact, a focus on a defendant's intentional commitments actually provides additional even stronger grounds to reject the hierarchy.

II. AVOIDANCE COMMITMENTS: RELUCTANCE RECONSIDERED

In Part I, I described two strategies for defending the PKRN hierarchy, both grounded in the doctrine of double effect, one by way of the purposeful and non-purposeful agents' quality of will as manifested by their PKRN mens rea states, one by way of the purposeful and non-purposeful agent's modes of reasoning, including intentional commitments which are distinct from their reasons for acting and constituted in part by their PKRN mens rea states.

If we look only at the weight that agents give the well-being of their victims as manifested in their practical reasoning, we lack a general account of why purposeful agents should be more liable than reckless agents. In fact, for reluctant purposeful homicides like those surveyed in the intimate partner homicide study by Browne and Williams, and callous reckless police homicides like that committed by Chauvin, it gives us a reason to treat police homicides as more liable than the intimate partner homicides.

---

[85] See *infra* Part IV. Doctrinal Implications and Avenues of Reform for more detailed discussion of why such doctrines fail to capture the diminished liability of the reluctant purposeful agent, and how such doctrines might be amended.

[86] *Id.*

A more promising defense of treating intentional wrongdoing as especially liable emerges only when we broaden our scope and look at the *intentional commitments* manifested by their actions. When we look at the defendants' respective intentional commitments, the stance of the reluctant purposeful agent toward their victim appears to be substantially worse than the callous reckless agent. Because the purposeful agent needs the victim to be harmed in order to achieve their goal, they will be disposed to continue making choices to harm the victim if their initial action fails to successfully produce the harm. The fact that the callous agent would not have similarly chosen to cause harm to the victim in those nearby counterfactuals where the victim was moved out of harm's way appears to show both that they are less dangerous and that they manifest less culpable failures in their modes of reasoning, than a purposeful agent, whether reluctant or not.

In this Part, I will argue that this 'tracking commitment' argument in support of the current mens rea hierarchy is still ultimately unsuccessful. I will show that the argument fails to consider equally or more problematic dispositions of the indifferent agent, and other more laudable dispositions of the reluctant purposeful agent. When these other dispositions are brought to light, we can see that a reasons-responsiveness account of culpability that includes a dispositional analysis of the agent's psychology actually undermines, rather than supports, the PKRN hierarchy.

To begin with, consider again the kinds of Trolley scenarios that defenders of the PKRN hierarchy employ to help reveal the problematic 'tracking dispositions' of the purposeful agent. Consider two agents with the same aim, getting to some goal, G, one of whom is completely indifferent to human life, one of whom finds the value of human life a strong reason to refrain from acting in ways which would result in the loss of such life (though not strong enough to outweigh the value of G). Suppose that in order to achieve G, the reluctant purposeful agent must push V onto the tracks, whereas the indifferent agent must drive through V, who happens to be on the tracks at the time.

The scenarios described are structurally identical to that of the traditional Trolley problem (where G stands in for the goal, in the original scenarios, of saving the lives of the five). And as we saw in Part I, a dispositional analysis of the two trolley scenarios may *seem* to favor increased liability for the reluctant purposeful agent. As the proponents of the doctrine of double effect note, though both agents may cause the same harms in the cases described above, there is a class of nearby counterfactuals where the callous agent and reluctant purposeful agent will diverge. If we move the victim out of the way of the tracks, or put a fence around the victim so that the tracks are no longer the most direct route to G, the callous agent will no longer be disposed to make choices that cause harm to V. This is precisely because they are callous. Since they don't care about V, one way or the other, they will not change their plans in order to hurt (or help) V. But the reluctant purposeful agent, despite valuing V's life, will still choose to harm V in this counterfactual. Since harming V is still a necessary means to achieving G, their desire for G will dispose them to make choices which 'track the harm' to the victim despite barriers we put in the way.

But it is a mistake to infer from this one set of counter-factual cases that the purposeful agent's 'modes of reasoning' will always guide them to 'track the harm' to the victim. In fact, this conclusion appears plausible only when we restrict our attention to an overly narrow set of counterfactual contexts, involving adding costs to disincentivize or deter problematic actors, made familiar to us from the Law & Economics literature.[87] But these are not the only available counterfactual scenarios to consider.[88]

Consider instead another category of counter-factual where we enable actors to do good, rather than disincentivize them from doing bad. Suppose that rather than moving the victim out of the way of the trolley, we open up an alternative means to G. Though more inconvenient, suppose the trolley has the option of taking some longer path around to G, that requires neither pushing V onto the tracks or driving through V.[89]

In this new variant of the trolley problem, unlike in the classic formulation, the relative dispositions of the two agents to harm the victim are reversed. Recall that the callous agent, because they don't care about V one way or the other, will simply take whatever means is easiest for achieving G. They will continue to act as they would have without the alternative path, even though their original route harms V. Since they don't see the fact that V is harmed as a cost, there will be nothing to motivate them to take on the extra inconveniences of the alternative means.

In contrast, the reluctant purposeful agent who values human life will be disposed to take on the inconvenience, and so take alternative means that don't involve causing V's death. If another alternative means to G becomes available that doesn't involve pushing V onto the tracks, they will take that alternative, even if more personally burdensome. Because they treat V's death as a reason not to act, they are pursuing V's death only because it was a *necessary* means to some much larger goal. If they treat V's death as a weighty reason, they will be disposed to take on even costly burdens to take other means to achieve their ends, provided those alternatives are available.

Agents who have higher concern for the well-being of others will exhibit what we can label *avoidance commitments*: they will be disposed to attend to alternative means to their goals that do not require harming the things and persons they value, and will be disposed to patterns of practical reasoning that lead to them choosing those alternative means, even taking on additional burdens to do so, when those means become available.

In contrast, agents who have no concern for the well-being of others will exhibit no such avoidance commitments. Even if the means which results in

---

[87] *See supra* note 82.

[88] In fact, (as I will argue in Part III), they are not even the most nearby counterfactuals. For the importance of 'nearby' counterfactuals for causal analysis, see DAVID LEWIS, COUNTERFACTUALS (2001).

[89] I first introduce this variant in Antill, *supra* note 27.

foreseeable harm is no longer necessary, because another alternative means is available, the indifferent reckless agent who fails to value human life would have no reason to adopt the more inconvenient means (or even equally convenient means) and so would not be disposed to change course.

Attending to these cases helps bring to light the intuitive sense in which the reluctant agent values the life of the victim more than the callous agent. The fact that the reluctant agent values the victim's life more than the callous agent in their practical reasoning may not prevent them from tracking the harm in 'fence-placing' scenarios. But it does have an important effect on their dispositions to engage in harmful activity in counter-factual situations where avoidance becomes possible. The reluctant agent, despite pursuing the taking of a human life as a means, has dispositions to respond better in such counter-factual cases. And those dispositions derive from the fact that the reluctant agent sees the taking of the life as intrinsically dis-valuable, and so is ready to take alternative means to avoid the impermissible result in ways that the indifferent agent is not.

Returning to our initial motivating cases of police-perpetrated homicide, with these new counterfactuals in mind, we can see that the callous police-perpetrated reckless or negligent homicide appears much more morally problematic than the intimate-partner homicide of the reluctant purposeful agent. Recall that a key reason for the female perpetrated intimate-partner homicides was that it was the only means that appeared available to agents to escape their entrapping life situation. Absent any infrastructure, like emergency shelters or legal avenues for divorce, the death of their spouse was the only perceived means available to them. But given the positive value they placed on human life, they would presumably have been disposed to take those alternative means if available, and so would not have engaged in harmful behavior toward the victim.

In contrast, part of what makes Derek Chauvin's actions seem so morally repugnant is that he was not just counterfactually disposed to ignore alternatives, he ignored the abundant *actual* alternatives he had actively available for effecting his goal of arresting George Floyd besides the fatal chokehold he chose to perform as a means of arrest. Indeed, it is the actual presence of these alternatives which shows how little concern for Floyd's life could have been manifested by Chauvin, consistent with his actions.

So far, I have been arguing for the culpability of a callous reckless or knowing agent who is aware of, and tolerates, a foreseen but unintended risk of harm to their victim. But one important upshot of the dispositional account I am proposing is that my diagnosis of the callous agent's culpability can also explain how a callous *negligent* agent might be guilty of the same kind of culpable indifference to the welfare of the victim as might a reckless or knowing agent.

That a negligent agent could be culpable of callous indifference to human life is less obvious on the 'quality of will account' from section I.A. Indeed, many proponents of such a view, like Alexander and Ferzan, who hold that the Model Penal Code system of criminal liability should be reformed to treat callous reckless agents as

more culpable, hold, for the very same reason, that negligent agents should not subject to any criminal liability at all.[90] This is because, on the Alexander and Ferzan account, agents are culpable to the degree that they actually failed to give appropriate weight to some reason in their practical deliberation. A reckless agent, for example, is subjectively aware of some risk of harm – a reason to refrain from acting – but assigns that risk of harm an inappropriately low weight in their practical reasoning. A negligent agent, however, who is not subjectively aware of any risk of harm, will therefore be incapable of treating the harm as a reason and so *a fortiori*, incapable of giving that reason insufficient weight.

When we look to an agent's broader commitments to be disposed to take proper account of others' well-being in their practical reasoning, however, things are even worse for the callous negligent agent than for the callous reckless agent. If the defendant fails to have sufficient concern for their victim to even factor into their deliberation the possibility of the victim being harmed in the first place, it is hard to see what would prompt them to search for, and so even become aware of, the alternative avenues or means that would avoid such harm, much less see what would incentivize them to take those alternative avenues when they are less convenient. So the callous negligent agent, who is negligent in virtue of failing to care enough about their victim to attend to whether their actions might affect the victim, will also fail to have avoidance commitments. In counterfactuals where they *were* aware that they would risk harm to the victim, they would have proceeded anyway. The same lack of concern which explains their negligence, and which their negligence manifests, also explains their disposition to behave wrongfully in avoidance counterfactuals.

In evaluating negligence, then, we must evaluate the question of *why* the agent was subjectively unaware of a risk of foreseeable harm. Negligence may be exculpatory if it were due to genuine forgetfulness, or some other morally neutral cognitive failing. But if the lack of awareness is the result of an uninquisitiveness about the risk to others, explained by a lack of concern for others' welfare, the lack of awareness is no excuse. And actions that demonstrate such a lack of awareness manifest the same lack of concern for others – or more – than would have been manifested if the agent had been aware and decided to proceed anyway.[91]

Return, for illustration, to the case of Derek Chauvin. The jury was convinced that given the evidence available to Chauvin – the fact that he had been sitting on Floyd's neck for over nine minutes, the warnings of the horrified bystanders, and the pleas of George Floyd himself – that he was subjectively aware there was a risk of death to Floyd, which he subsequently accepted in choosing to continue to act as he did.

---

[90] *See* Larry Alexander & Kimberly Ferzan, *Against Negligence Liability in* CRIMINAL LAW CONVERSATIONS 272-294 (Paul H. Robinson et al, eds., 2011). *Compare* Shiffrin, *supra* Note 83; A. P. SIMESTER, FUNDAMENTALS OF CRIMINAL LAW: RESPONSIBILITY, CULPABILITY, AND WRONGDOING 237-261 (2021).

[91] This is why, for example, in the film *Casablanca,* Rick's reply to Urgate's claim that Rick despised him: "I probably would, if I gave you any thought at all" is so devastating. *Casablanca* (Warner Bros. 1942).

But suppose, as Chauvin claimed, that he really was not subjectively aware of a risk to Floyd's life. In this case, such a lack of awareness would require Chauvin to have treated Floyd's life as mattering so little that he did not even bother to stop and consider whether his actions might risk killing Floyd; or that he thought so little of Floyd, and so highly of his own experience, that he gave no evidential weight to the pleas of Floyd, the crowd, and his fellow officers.

In this case, Chauvin's mistake of fact, which led to him being negligent, rather than reckless, would still have been part of a broader lack of concern for George Floyd. If Chauvin were reckless, it manifested in his ignoring the risk to Floyd's life; if he were negligent, it manifested in his ignoring the very *question* of whether there was such a risk. In either case – whether reckless or negligent – the behavior manifests a near total lack of concern for the well-being of his victim, and so a culpable commitment to dispositions of practical reasoning where the value of that life would be discounted – whether or not Chauvin had been aware that the value was at play.

Although the difference between whether Chauvin was reckless or negligent has important legal consequences for his criminal liability on the current PKRN mens rea hierarchy, an investigation of his intentional commitments helps explain why from the perspective of normative jurisprudence, the question of whether he was, as a result of his callousness, negligent or reckless toward Floyd's death, should be irrelevant.

So far, the discussion of avoidance commitments has centered on the twin guiding cases of non-purposeful police perpetrated homicides and reluctant purposeful homicides by survivor defendants. But the presence of avoidance commitments can be manifested in defendants with far less sympathetic motives, and against more sympathetic victims. Nonetheless, the presence (or absence) of avoidance commitments frequently does a better job than the traditional PKRN hierarchy in sorting both the least and most culpable defendants, and the least and most dangerous defendants, even in these more commonplace cases of criminal wrongdoing.

Consider a pair of more quotidian homicides, driven by pecuniary motives. Imagine, for example, a defendant, D1, who causes the death of their victim by burning down a house while aware that the victim is inside, as a means of getting $10,000 of life insurance money. Compare this to a defendant, D2, who burns down the house while aware that the victim is inside, as a means of getting $10,000 of home insurance money.

This kind of hypothetical is commonly used in criminal law texts to introduce the difference between the mens rea states of purpose and knowledge. In the case of D2, the death of the victim is a foreseen, but unintended consequence of their plan to burn down the house for the home insurance money. D2 does not require V's death as a means of achieving their further pecuniary objective, and so the death is not their "conscious object." In the case of D1, on the other hand, the killing of the victim is a necessary step in their plan to get the life insurance money, and so purposeful.

With slightly more variation, we can create reckless or negligent versions of the same case by varying the defendant's subjective awareness of the degree of risk

that the victim is inside the home. We can imagine a reckless defendant who burns down the house as a means of getting the $10,000 of life insurance money while aware of some probability that the defendant may be in the house (and so "consciously disregards a substantial and unjustifiable risk" of harm to the victim). And we can imagine a negligent defendant who burns down the house without even considering whether someone might be inside, and so acts in the face of a foreseeable "substantial and unjustifiable risk" of which "he should be aware" but is not.

Thinking about this kind of case helps distinguish the account of culpability I am putting forward here from both a reasons-responsiveness account of culpability underling the PKRN hierarchy, and from the kind of quality of will account underlying Alexander and Ferzan's proposed alternative to the model penal code according to which a defendant's mens rea should be a function of a defendant's reasons or motives for action. It will also show how neither account of culpability is particularly well suited to capturing the kind of mitigating factor of reluctance which, on my view, ought to have a central place in our culpability calculations.

On the reasons responsiveness account of culpability which underlies the PKRN hierarchy, the fact that the first defendant killed purposefully marks an important difference from the other cases. Because the first defendant manifests an intentional commitment to track the harm to the victim, the reasons-responsiveness account holds that they are both more dangerous (they will continue trying to kill the victim if they escape, they will try to light the fire in a way that maximizes the victim's probability of death, etc.) and more culpable than the knowing or reckless agent who simply tolerates the known (or probable) death of the victim.

In contrast, on the account I am putting forward in terms of avoidance commitments, what matters most is not whether the various defendants acted purposefully (or knowingly, recklessly, or negligently), but the degree to which they were committed to avoiding criminal wrongdoing. And for any of the four cases of purposeful, knowing, reckless, or negligent defendants, we can imagine a version with or without such avoidance commitments.

With respect to purpose, for example, we can construct a version of the hypothetical where the purposeful agent manifests avoidance commitments, and a case where the purposeful agent lacks such commitments. Suppose, for example, that D1, the purposeful defendant, desperate for money, commits the homicide only after trying, and failing, to find any other lawful avenue to make the ten thousand dollars. But had they found any legal method of earning the money, even if it was through menial or uncomfortable labor, they would have eagerly taken that alternative rather than pursue the means of killing V for the insurance money. Though they were willing to kill for 10,000 dollars, they choose killing as their means to acquiring the money only as a last resort. Contrast this case with a reluctant purposeful defendant who lacks avoidance commitments. Consider, for example, D1-Prime, the professional assassin who is not desperate for the money, who has plenty of alternative ways of getting the thousand dollars, but who enjoys the challenge and flexible hours of their chosen profession.

Similarly, for a reckless defendant, like D3, we can construct a case where the reckless agent has (or lacks) avoidance commitments. Constructing a case of a reckless agent who lacks avoidance commitments is relatively straightforward. Even for reckless defendants who lacked any method to get the money other than burning down the house for the insurance money, it is easy to imagine plenty of other alternative less harmful ways of pursuing that means, such as simply waiting until they were certain the house was empty before lighting the fire. Suppose that despite these alternatives, impatient, the reckless defendant simply lights the fire when it is most convenient, while aware of the likelihood that the victim was in the house. Such a defendant, like Chauvin, would be a reckless defendant who lacks avoidance commitments.

It is a little more difficult to construct a version of the reckless commission of the crime where the defendant *does* manifest avoidance dispositions. Still, though more unusual, such cases can be imagined. We can contrast the callous reckless agent with a reluctant reckless agent who suspects the victim is inside, but (let us suppose) has only one chance to burn down the building for the insurance money, though had any alternative opportunity to burn down the building been available that risked less chance of harm to the victim, they would have been strongly motivated to take that alternative method instead.

Though issues of self-defense, necessity, victim-provocation, and extreme emotional distress and other forms of reduced responsibility are not implicated in the arson hypothetical, D1 still shares with survivor defendants the existence of avoidance commitments. And though issues of abuse of police power, racism, or civil rights violations are not implicated, D3 still shares with Chauvin an absence of avoidance commitments. And still, in these cases, D1 appears much less culpable, and much less dangerous, than D3. The two reluctant agents who are committed to avoiding harm, D1 (the reluctant purposeful defendant) and D3' (the reluctant reckless defendant), have more morally in common with one another than the two callous agents, D1' (the callous purposeful assassin) and D3 (the impatient reckless arsonist). Despite both D1 and D1' purposefully killing V, and despite both being motivated to do by the ten-thousand-dollar payout, the absence (or presence) of avoidance commitments intuitively makes the two cases quite different. And despite the fact that the impatient callous agent who declines to wait until they know the house is empty is reckless, not purposeful, their lack of avoidance commitments makes D1' and D3 importantly alike.[92]

---

[92] These pairs of reckless and purposeful agents are more similarly situated than the cases of Chauvin and Survivor-defendants. Here, they have engaged in the very same act, with the very same result, and out of the very same pecuniary motive. Still, it's true that in adding background to the defendants to change their avoidance dispositions, I have changed more about the defendant than whether the defendant was purposeful, knowing, reckless, or negligent. This may seem dialectically problematic. If the doctrine of double effect ("DDE") were claiming simply that, ceteris paribus, intentional agents are more culpable than knowing agents, and we then change not just whether the agent is intentional or knowing, but also other factors about the agent's background, it may appear that all else is no longer equal. *See, e.g.,* Sarch *supra* note 31, at 460-461. Whether it is dialectically appropriate depends on what question we are pursuing. Alex Sarch, for example, argues for a version of the DDE according to which intentional harms are *pro tanto* more culpable than the harms committed by knowing agents. And because of the different reasons of the two actors, it's not clear that the cases discussed in this section constitute a counter-example to such a claim. However, Sarch also claims that his "main aim is . . . . to

As this second pair of cases begins to suggest (a suggestion I will draw out further in Part III), the failures of the current mens rea regime to take reluctance into account are not limited to these exceptional cases. The core underlying problem with the standard PKRN hierarchy is that it fails to distinguish between the purposeful agent who commits a crime reluctantly out of poverty, abuse, or other forms of desperation, and the callous purposeful agent with plenty of alternatives who commits crime casually whenever it happens to be in their self-interest. The basic motivating idea of the Article is that despite their different place in the traditional PKRN mens rea hierarchy, the two defendants who are reluctant have more in common than the two who are not.

This second pair of cases also helps sharpen the difference between the account of avoidance commitments being proposed here from alternative quality of will accounts, like that of Ferzan and Alexander. A 'quality of will' account of culpability focused on the agent's reasons for acting is not especially well-suited to capturing intentional commitments, like avoidance commitments. Like tracking commitments, avoidance commitments can vary across agents acting with identical motives. And so, looking at two defendants who both choose to cause harm for some monetary benefit (e.g., killing a victim for 10,000 of insurance money) will not distinguish between the defendant who made this choice only as a last resort, and the victim who made the choice despite having available alternatives.[93]

To be clear, the claim is not that other features make no difference to culpability. Or even that purpose makes no difference to culpability. A Chauvin-prime who not only lacked avoidance commitments with respect to Floyd's suffering, but who actively sought out Floyd's death, would of course be more culpable and more dangerous than the actual Chauvin. The claim, instead, is that the presence or absence of avoidance dispositions is a feature of the situation with a great deal of importance to culpability, and which the law, due to its emphasis on intentional wrongdoing, is unable to recognize through either its mens rea states or its affirmative defenses.

---

defend a version of the DDE that can do what is needed for the purposes of justifying the criminal law." *Id.* This is why, for example, Sarch restricts his version of the doctrine of double effect to cases of unjustified action. As he notes, while this version of the doctrine may not "give all proponents of DDE everything they want, it would still give criminal law theorists everything they need." *Id.* at 460-461. But what criminal law needs is that a purposeful agent is more culpable than a knowing agent holding fixed all the other elements of the crime relevant to criminal liability. Thus, we must hold fixed the material elements, including the harmful consequences, as well as any relevant psychological facts that could underlie an affirmative defense of excuse or justification. But this is not the same thing as holding fixed every other aspect of the agent's psychology, because some of those aspects (most crucially, the agent's reasons for acting, when those reasons do not suffice to justify the action) do not affect degree of criminal liability. Since differing motives among reckless and purposeful agents do not affect those agents' respective criminal grades, a defense of the criminal law's mens rea regime needs to show that purposeful agents are more culpable than reckless agents who commit the same criminal offense, *regardless* of those differing motives. And it is precisely this which these examples seek to show cannot be done.

[93] *But see* Christopher Lewis, *Inequality, Incentives, Criminality, & Blame,* 22 L. THEORY 153 (2016) (arguing that, due to diminishing marginal returns, the very same monetary reward may constitute a reason with different force, or utility, for agents with different resources).

This claim has stronger and weaker flavors as applied to these homicide cases (and to other criminal offenses more broadly). The strong version of the claim is that avoidance commitments are one of the *most important* of these factors for culpability. And so, to the extent that criminal law seeks to grade crimes according to culpability, or set minimum mens rea thresholds for an offense, avoidance commitments are a better measure than the current mens rea states like purpose or degree of subjective awareness, when the two come apart. If we're looking for the more fundamental cuts in culpability, the mens rea states that should set the boundaries between murder and manslaughter, or between first-degree murder and second-degree murder, the argument is that avoidance commitments play a distinct and more fundamental role in culpability than the tracking commitments constitutive of purpose. In the case of the arsonists, for example, it may be that the reluctant purposeful arsonist is more culpable than the reluctant reckless arsonist. But the two reluctant arsonists have much more morally in common than the two callous defendants (the reluctant purposeful arsonist and the professional assassin). And so, avoidance commitments ought to be used to make the fundamental cuts, rather than the mens rea categories of the current PKRN hierarchy.

A weaker claim is just that avoidance dispositions should play *some* important role in culpability assessments in the law. Even if we don't make avoidance commitments a core part of the mens rea hierarchy, the evidence that a survivor-defendant went to extraordinary lengths try to find another way to escape her situation before killing her spouse, ought to provide *some* avenue for decreased criminal liability, even if it takes the form of an imperfect affirmative defense, or a special mitigating mens rea state, which we might append to the PKRN hierarchy along with already existing mitigating states like passion-provocation or extreme-emotional disturbance.

While I will attempt in Part III to motivate the stronger version of the claim, either will suffice for the purposes of the paper more generally in its argument for the need for doctrinal reform. The problem is not only that avoidance dispositions fail to play a *central role* in culpability assessments. It's that the current code, with its emphasis on purpose, and the way purpose has shaped both the mens rea hierarchy and defenses, fails to make space for *any* straightforward claim for mitigation on such grounds. As I will argue in Parts IV and V, attending to the importance of avoidance commitments in our mens rea doctrine is an essential step in recognizing the reduced culpability of many defendants who engage in criminal activity intentionally, but as a last resort, due to factors like poverty, social disadvantage, or structural racism.

III. AVOIDANCE COMMITMENTS, POWER RELATIONS, AND ENVIRONMENTAL DEPRIVATION

We can now see that defenders of the PKRN hierarchy were too quick in assuming that the dispositions of the purposeful agent who aims at, and so tracks, the harm to others are necessarily worse than the dispositions of a callous agent. It turns out that it depends which counterfactuals we are attending to.

If we look at the counterfactual dispositions in 'tracking cases' where we remove the victim from the harmful consequences of the agent's plans, either by

'building fences' to protect the victim by making the action more difficult or by 'moving the victim out of harm's way,' the dispositions of a purposeful agent will be worse. So long as harming the victim remains a necessary means to their ultimate goal, they will continue trying to harm the victim.

However, if we look at 'avoidance cases' where the harmful action is no longer a necessary means to the agents' goals – either by removing barriers that prevented the agent from having any alternative means available, or by working to build such alternative paths, the callous agent's dispositions will be worse. Because they do not care about the victim, they will not cease the harmful behavior unless provided some further deterrence.

This is not (yet) a decisive argument for abandoning the traditional PKRN mens rea hierarchy. But it does show, at a minimum, that the traditional mens rea hierarchy is less innocuous than it seems. It is not simply entailed by the premise that criminal liability should mirror the quality of an agent's intentional commitments as manifested in their actions. Instead – whether implicitly or explicitly – the law's attitude toward these two kinds of agents reflects a prioritization of building walls over building new paths or removing legal barriers to existing paths.

I will now argue in Part III that this prioritization is a mistake. The law is mistaken in the principle that it is easier to build fences around victims than it is to open up new opportunities for offenders. In fact, there are many cases where it is much easier for society to create new paths or remove legal barriers, and so allow reluctant purposeful offenders an offramp to avoid criminal behavior, than it is to build the kind of protective policies necessary to deter callous agents from thoughtlessly harming those unlucky enough to be caught in their way.

This argument will proceed in two steps. I will begin by continuing the discussion of the twin cases of police-perpetrated homicides and survivor-defendant homicides, and show how these cases challenge the State's commitment, implicit in the current mens rea hierarchy, to deterrence through negative disincentives. I will then show how these lessons generalize in ways that should motivate us to fundamentally rethink criminal law's emphasis on purposeful wrongdoing more generally, by showing how the avoidance commitments manifested by survivor defendants who commit purposeful homicides are likely widespread among criminal defendants engaged in purposeful wrongdoing.

Police homicides are representative of the difficulty of deterring defendants who lack avoidance commitments, even when their wrongdoing is not purposeful. There are, in typical police encounters, plenty of available avenues by which police might be able to achieve their ends of preserving order that involve less endangerment to those being arrested. Social scientists have documented a variety of policing reforms including de-escalation policing, procedural justice policing, and trauma informed policing which are designed to 'slow down' encounters and avoid direct confrontations

which may foreseeably lead to more police violence.[94] Increasingly, these methods are trickling down into policy proposals and police trainings.[95]

This movement is well illustrated in the recent federal CCJ Task Force's "Path to Progress Report." The Task Force recommends a move away from the dominant current policing method which adopts a "militaristic warrior model, employs a stress-based approach, and emphasizes intensive physical demands, firearms proficiency, psychological pressure, and enforcement rather than trust building and problem solving."[96] In its place, the task force explains, the evidence suggests we must move toward a new model which places a "far greater focus on communication and critical thinking skills, social interaction and de-escalation tactics, and principles of procedural justice."[97]

However, the adoption of these methods has faced stiff resistance. Social scientists have isolated, as a primary factor, high levels of Social Dominance Orientation (SDO) among those police officers most likely to engage in police misconduct.[98] Social Dominance Orientation (SDO) is a personality trait indicating a general preference for hierarchical structures and for the social dominance of in-groups over out-groups.[99] Individuals with high levels of SDO are more likely to engage in discriminatory behavior against outgroups, to believe that coercion is necessary, to resist procedural justice, and to support or engage in use of force against outgroups. In short, individuals with high levels of SDO are less likely to believe that the lives of those in the out-group matter. Research suggests that the current warrior model of policing helps inculcate such an orientation, so that "SDO tends to be higher in police officers compared with members of the general public, college students, and public defenders, even controlling for demographic variables"[100] and that these high rates of SDO can help explain the "increased use of force in Black and Latino neighborhoods" by police, which cannot be accounted for by controlling for differences in crime- and poverty- rates.[101]

The fact that "individuals with higher levels of SDO tended to choose to go into policing"[102] suggests that police who engage in misconduct will often lack the avoidance commitments described in the previous section. Individuals with higher levels of SDO, who have less concern for the interests of those they police, will not be motivated in their practical reasoning to seek out, or adopt, alternative methods

---

[94] Emily Owens et al., *Can you Build a Better Cop? Experimental Evidence on Supervision, Training, and Policing in the Community,* 17 CRIMINOLOGY AND PUBLIC POLICY 41, 41 (2018).
[95] CCJ Task Force On Policing, "The Path to Progress: Five Priorities For Police Reform 1 (May 2021); The President's Task Force on 21st Century Policing, Final Report (May 2015).
[96] *Id.* at 3.
[97] *Id.*
[98] Jillian K. Swencionis and Phillip Atiba Goff, *The Psychological Science of Racial Bias and Policing*, 23 PSYCH., PUB. POL'Y, & L 398, 399 (2017).
[99] *Id.*
[100] *Id.*
[101] Phillip Atiba Goff, *Identity Traps: How to Think About Race & Policing*, 2 BEHAVIORAL SCIENCE & POL'Y 11, 12 (2016).
[102] Swencionis & Goff, *supra* note 98*,* at 403.

which are less efficient means toward their ends but which avoid the unintended side-effect of harming others.

This failure of concern by officers with high levels of SDO for the lives of those they police – and, in particular, a failure to appreciate the way Black lives matter – can explain why it is frequently so difficult to implement specific policing reforms in the field without first reforming the underlying problems in police culture.[103] The problem is that, absent some disposition to care about the well-being of the population being policed, it can be very difficult to convince or motivate police officers who are disposed to engage in misconduct to take those routes instead of the simpler more familiar policies of chokeholds and other more traditional 'dominance-and-control' methods of policing with much higher risks.

And, as we have seen so painfully over the last decade, it is much more difficult than the law presupposes to 'build fences' around members of policed populations to deter a callous or indifferent police officer from pursuing the 'path of least resistance' even if this path involves higher risks of harm to those policed. Various attempts to do so with body cameras, police review boards, and other populace protections have had limited success, at best.[104]

In contrast, with political will and concern, it is often comparatively easy to create new alternatives for those reluctant purposeful criminals whose violence is the result of desperate circumstances. Take the population of female perpetrated intimate partner homicides from Browne & Williams' study.[105] In those cases, many of the homicides were the result of cruel and draconian divorce laws which created legal barriers to those defendants escaping their situation by anything other than violence. In the intervening years, these laws have been changed. When those barriers were reduced, those States saw a dramatic decrease in violent crime among that population.[106]

What this comparison of the two cases of non-purposeful police-perpetrated homicides and purposeful survivor-defendant homicides teaches us is that it can frequently be far easier to reduce crime by enabling reluctant defendants engaged in purposeful wrongdoing to do good instead, than it is to disincentivize callous non-purposeful wrongdoers from acting in ways that cause collateral harms to others.

But how significant a problem this is for the criminal law's mens rea hierarchy, and its general emphasis on purposeful wrongdoing, depends on whether the class of reluctant purposeful actors is limited to the relatively rare case of female-perpetrated

---

[103] *See* Goff, *supra* note 101.

[104] *See, e.g.,* Barak Ariel et al., *Wearing Body Camera's Increases Assaults Against Officers and Does Not Reduce Police Use of Force: Results From a Global Multi-Site Experiment,* 13 EUROPEAN JOURNAL ON CRIMINOLOGY 774 (2016).

[105] Browne and Williams, *supra* note 9.

[106] *Id. at* 75 (providing statistical analysis indicating that "the availability of [legal and extralegal resources for abused women] is associated with a *decline* in the rates of female-, but not male- perpetrated homicides."); *See also, e.g.,* Betsy Stevenson & Justin Wolfers, *Bargaining in the Shadow of the Law: Divorce Laws and Family Distress,* 121 QUARTERLY JOURNAL OF ECONOMICS, 267 (2006).

homicide, or whether it generalizes to more cases of homicide, and purposeful crime more broadly.

It is, of course, ultimately an empirical question what proportion of purposeful crime involves reluctant purposeful agents, and what proportion of reckless/negligent crime involves especially callous agents. This question is difficult to answer definitively at present. Since avoidance commitments have not received the same scholarly attention as tracking commitments, we lack empirical studies investigating their prevalence directly. One hope of the present Article is that by showing the potential importance of avoidance dispositions, it might spark more empirical research into the question of their prevalence.[107]

Still, in this section, I will attempt to make good on the claim that, given the evidence we *do* have, we should expect (at least provisionally) cases of purposeful wrongdoing where the defendant manifests avoidance dispositions to be widespread.

The argument for this claim is partly theoretical, and partly empirical. I will provide a theoretical argument that given the general nature of power relations, we should expect the empirical results to show disproportionately more reluctance among purposeful wrongdoers than reckless or negligent wrongdoers. We should expect purposeful crime to frequently be the result of an absence of viable alternatives due to social disadvantage, while, in contrast, we should expect those with wealth and resources who lack avoidance dispositions to be proportionately more likely to engage in reckless or negligent wrongdoing. I will supplement this argument with empirical evidence, when available, the majority of which tends to confirm this prior expectation.

We can begin by returning to the observation that it is possible, at least in principle, to construct more 'run of the mill' cases of purposeful homicide which also exhibit avoidance dispositions. The defendant who purposefully kills their victim for the life-insurance money may be a hardened, callous professional killer who acts without compunction. But they are more likely to be a defendant who is in desperate economic circumstances who would have chosen to pursue alternative means (and perhaps in fact tried and failed to pursue such means) to make a living, had they been available.[108] Contrast such cases with a reckless arsonist who kills a victim as a collateral consequence of burning down a building for the home-insurance money. Since the victim's death was a side-effect, rather than a necessary means of getting the insurance money, the idea that no safer alternatives were available is less plausible. It is possible that such a reckless defendant engaged in arson only as a last resort, and had only one opportunity to burn down the building and so had no alternatives that involved less risk to the victim. But absent such a story, it is difficult to see how the reckless arsonist could have lacked alternatives (like simply waiting until they were sure the building was empty). Like the reckless nightclub owner who bars the fire exits to prevent patrons skipping the entrance fee in order to increase his profit margin, the impatient reckless

---

[107] I plan on contributing to this research in future work.

[108] *See, e.g.,* Jerel Ezell, *Understanding the Situational Contexts for Interpersonal Violence: A Review of Individual-Level Attitudes, Attributions, and Triggers in* 22 TRAUMA, VIOLENCE, AND ABUSE 571 (2019).

arsonist who causes the death of his victim will have foregone other less harmful alternative means to achieve material security, and so lack avoidance dispositions.[109]

A similar analysis can apply more broadly beyond homicide. In the case of property crimes, for example, many cases of purposeful theft are the result of various social or economic pressures that make crime appear to be the only viable option for dispossessed populations.[110] By creating new programs to provide alternative legal paths to economic and social self-sufficiency, we could make it so that crime was no longer the only means available for achieving the offender's goals.[111] Indeed, given that reluctant purposeful offenders place positive value on the well-being of others, the reluctant purposeful agent would presumably take these other alternatives that avoid harm to others, even if they led to *less* benefit than the criminal alternatives, so long as they would still allow for minimally tolerable material circumstances. [112]

In contrast, there is no easy way to ensure the callous reckless or negligent agent will be similarly incentivized. We can ensure that callous landlords or white-collar criminals do not feel that they must commit a crime in order to enjoy minimal levels of economic and social security (indeed, such non-criminal alternative means to material security already exist for such defendants). But it is much harder to ensure that there will *never* be situations where such white-collar criminals might be a little more materially well-off by taking actions that risk some physical or financial harm to others, and so no easy way of disincentivizing them from causing harm if they lack any practical commitment to take such non-harmful alternative means to their ultimate goals, when such means are available.

And yet, because the mens rea standard for white collar crime is typically knowledge or purpose, callous white collar defendants who engage in reckless or negligent harm to the victims will likely avoid criminal liability altogether.[113] Whereas those who commit crimes of desperation, because those crimes are often purposeful, will be subject to the highest levels of criminal liability which purposeful wrongdoing, in the traditional mens rea hierarchy, entails.

---

[109] *See e.g. Commonwealth v. Welansky*, 55 N.E.2d 902 (1944).

[110] *See* TOMMY SHELBY, DARK GHETTOS (2018); Christopher Lewis, *supra* Note 93.

[111] This is a common observation by scholars and advocates of prison abolition and decarceral reform. Those in such movements frequently observe that the money spent on prisons could be better utilized by such societal programs. Proponents of 'transformative justice' approaches to criminal law typically focus on crime prevention, rather than culpability. Indeed, such projects are often seen as incompatible with retributivist or expressivist approaches to criminal law focused on the defendant's responsibility for harm. SEE, E.G., TOMMIE SHELBY, THE IDEA OF PRISON ABOLITION (2022). One goal of this Article is to show how these considerations, which are frequently utilized as grounds for external critique of the criminal justice system, also have force, even from the internal perspective concerned with the goals of culpability and deterrence.

[112] *See* Ezell *supra* note 108.

[113] See, e.g., Benjamin Levin, *Mens Rea Reform and Its Discontents,* 109 J. OF CRIM. L. & CRIMINOLOGY 491, N. 135 and accompanying text (2019).

This tendency for reluctance in purposeful defendants and an absence of reluctance in reckless defendants is of course not universal.[114] In the case of homicide, for example, there will be purposeful agents, like the contract killer of law school hypotheticals, who purposefully causes the death of their victim as a means of making more money even though they have plenty of money already to achieve material security and other available means of making a living.

Far more importantly, there are going to be cases of crimes of poverty or desperation that are performed recklessly. Cases of drug-based homicides are a prominent example. A defendant may recklessly cause the death of the victim by selling them drugs, knowing there is some probability that the drugs will cause the victim to overdose. However, they may be engaging in such criminal activity only because they lack any other means of making a living. (I discuss in Part V how the paper's proposed mens rea reforms can avoid heaping more criminal liability onto such defendants, and prevent the risk of their being classified as cases of depraved-heart recklessness, as they too often are in the current system).

Still, the idea that purposeful wrongdoing will be proportionately more likely to be reluctant (and will extend far beyond the case of survivor-defendants) shouldn't be surprising. It is what we should expect as a natural consequence of a close conceptual connection between power, means, and opportunities to avoid criminal activity.

It may be helpful to think about a distinction of Arthur Ripstein, made in the context of private law theory, between two categories of wrongs.[115] One kind of wrong is committed when someone uses another person (or other people's property) as a means toward pursuing their goals. In the context of criminal law, these are intentional or purposeful wrongs. They are cases where the wrong is the conscious object of a defendant's plan, rather than foreseen or foreseeable collateral side effect of that plan.

A different kind of wrong is committed by those who have their own means to pursue their goals (and so don't need to use other people's means) but who, when using their own means, fail to exercise sufficient care that they do not cause harm to other bystanders or third parties along the way. In the context of criminal law, these are wrongs of negligence or recklessness.

Those with more power and wealth have more ways of pursuing their goals. As Ripstein observes, they are, literally, people of means.[116] And all else being equal, the more means an agent has at their own disposal to pursue their goals, the less likely it will be that some criminal course of action is the only or necessary means to achieving that goal (even if it happens to be the most convenient means to their goal in any given context). So there will be fewer contexts where people of means will need to purposefully wrong another person by using them (or their property) as the only

---

[114] (which is way, as will be discussed in Part IV, I suggest we need to not invert the hierarchy, but replace/modify it to focus on dispositions)

[115] *See* RIPSTEIN, *supra* note 19.

[116] *Id.*

means to their desired ends. And, because they have more means of their own, there will be more contexts where they might exercise those means in various ways to achieve their goals, with more or less efficiency, and more or less risk of harm to other people. The result is that people of means (people in positions of wealth or power) will (a) be proportionately more likely to engage in reckless or negligent wrongdoing than purposeful wrongdoing and (b) more likely to lack avoidance dispositions when they do engage in such negligent or reckless wrongdoing. Since they have more ways of achieving their goal, people of means who *have* avoidance commitments will be able to act on their commitments by taking the alternative means they have available to them in the actual world, and so act with care in ways that successfully avoid risking known harming others. Conversely, those with means who *do* choose to risk harm to others will more frequently be those who lacked the commitments to utilize the means they have to seek other alternatives (since, if they had such commitments, they would have been disposed to take those alternative means and would not have engaged in wrongdoing). In other words, those with means who are committed to not harming others, won't harm others.

Meanwhile, the fewer means a person has, the more likely they are to lack legitimate means to their goals, and so the more likely they are to need to use other people (or their property) to increase the means at their disposal to achieve their ends. The result is that people who lack means will (a) be proportionately more likely to engage in purposeful wrongdoing and (b) more likely to manifest avoidance commitments when engaged in that purposeful wrongdoing. Unlike those with wealth or power who can act effectively on their commitment to avoid harming others, those who are committed to avoiding harm, but lack legitimate means to realize that commitment, will often be unable to manifest their concern for others by avoiding harm. Their commitment will be masked by the absence of any alternative means in their environment to avail themselves of.[117]And because those who lack means of their own will be more likely to need to use the other person to expand their means, their harm will more frequently be purposeful, and so subject to more liability on the current hierarchy.

Attending to avoidance commitments, then, can provide us with new insight into how the current mens rea regime functions to penalize socially disadvantaged defendants. It is not simply that the law fails to recognize environmental deprivation by excluding such factors in the federal sentencing guidelines, or as possible grounds for affirmative defenses like duress.[118] The PKRN hierarchy itself, in emphasizing disincentivizing wrongdoing over enabling right-doing, functions to further penalize the kinds of criminal activity we should most expect to be associated with social disadvantage, while mitigating the liability for the kinds of criminal activity we should most expect to be associated with abuses of power.

---

[117] A point I explore in more detail elsewhere in Antill, *Agency, Akrasia, and the Normative Environment*, 5 JOURNAL OF THE AMERICAN PHILOSOPHICAL ASSOCIATION 321 (2019).

[118] See, e.g., Miriam S. Gohara, *In Defense of the Injured: How Trauma Informed Criminal Defense Can Reform Sentencing,* 45 AM. J. CRIM. L. 1, 25-31 (2018); Joshua Dressler, *Exegesis of the Law of Duress: Justifying the Excuse and the Search for Its Proper Limits,* 62 S. CAL. L. REV 1331 (1988).

In contrast, as I will hope to show in more detail in the final two parts of the article, a mens rea regime that gave a more central place to avoidance dispositions can help the law take environmental circumstances into account in new and productive ways. As political philosopher Tommie Shelby has noted, environmental deprivation, even in those limited situations where it is allowed to be taken into account (as in cases of sentencing for capital offences), is usually thought to be relevant to culpability by showing that defendants who suffered from environmental deprivation have diminished responsibility for their actions.[119] This has two potential flaws. First, as Shelby notes, it threatens to dehumanize those defendants from socially marginalized populations, by denying their agency.[120] Second, in a system like criminal law, committed to a compatibilist conception of responsibility, treating environmental deprivations as diminishing an agent's responsibility may appear to threaten the entire edifice of culpability assessments more generally.[121] It may appear that we cannot allow mitigation for diminished responsibility due to environmental deprivation without allowing mitigation for *every* defendant, since every defendant's will is shaped by their environment, regardless of deprivation. This likely explains, in part, why the criminal law is so often resistant to treating environmental factors as exculpatory.[122]

Attending to avoidance dispositions lets us sidestep this debate. It allows us to control for some of the objectionable moral luck involved in a defendant's environment which the current system struggles to do, without requiring that we abandon the practice of using culpability to help apportion criminal liability altogether.[123]

Ignoring a defendant's environmental circumstances when determining degree of culpability or blameworthiness introduced a potentially problematic form of moral luck in the sense that many of us would have been committed similar crime in similar counterfactual circumstances; and many defendants with poor environmental circumstances would not have committed crime had they been in our counterfactual circumstances.

Using avoidance commitments as a mens rea state allows us to focus on the underlying psychological states (an intentional commitment to avoiding harm) which explains those counterfactual dispositions. Giving avoidance commitments a central place in our theory of culpability thus allows us to control for some moral luck without denying that the defendants are still fully responsible for their wrongdoing. To say someone has avoidance commitments is not to say that they are not answerable for their choices, or even to deny that their choices manifest a culpable state of mind. It

---

119 *See* SHELBY *supra* note 110; Dressler, *supra* note 118.

120 *Id.*

121 *See, e.g.,* Stephen J. Morse, *Compatibilist Criminal Law, in* THE FUTURE OF PUNISHMENT *(*Thomas A. Nadelhoffer ed., 2013)

122 *Id.*

123 I suspect that some moral luck is unavoidable in a workable system of criminal law. *See* Sandy Kadish, *Criminal Law and the Luck of the Draw,* 84 J. OF CRIM. L. & CRIMINOLOGY 679 (1994). However, to the degree such moral luck can be minimized without any further cost to the system, this should be an attractive desideratum of a mens rea regime.

is simply that they manifest a less culpable state of mind than someone with more options who engages in the same criminal activity.

Of course, not everyone who commits purposeful crime after having the misfortune to suffer from environmental deprivations or victimization will fit the mold of the reluctant purposeful agent with avoidance commitments described above. One of the cruelties of suffering victimization or growing up in poverty is that these formative circumstances can also harden one's heart, so that, as a result of environmental circumstances, an agent engages in purposeful crime due to insufficient concern for other people. At the extreme, such poor formative circumstances can produce cases like Robert Harris, the serial killer who developed psychopathy as a result of extreme abuse suffered as a child.[124]

Because they lack avoidance commitments, my proposal will not provide an avenue for reducing the criminal liability of such defendants. In Thomas Nagel's terms, avoidance dispositions help control for the ways in which environmental deprivation can contribute to 'circumstance luck' but not to 'character' luck.[125] For a defendant who is concerned about others but, due to environmental factors, is driven to purposefully engage in crime as a last resort, avoidance commitments can explain why they are both fully responsible for that action and yet less culpable. However, if a defendant, due to environmental deprivation, grew up to simply lack concern for other people, and so engaged in criminal activity that they would not have engaged in but for that environmental deprivation, this account would do nothing to reduce their liability.

Whether this is a virtue or a limitation depends on the degree of one's commitment to compatibilism. It is a deep and difficult problem how the criminal law ought to treat such defendants.[126] On the one hand, such defendants demonstrate a lack of concern for others which is at once both extremely dangerous and apparently culpable. On the one hand, it may appear unfair to punish such defendants when they did not get to choose the kind of person they became (especially when the state doing the punishing is complicit in causing or permitting the poor formative circumstances).

For those opposed to allowing poverty or environmental circumstances to mitigate because of a commitment to the position that defendants like Robert Harris should not be given reduced criminal liability, avoidance commitments provide a way of demarcating a subset of those suffering from environmental deprivation who are uncontroversially less culpable and less dangerous, and granting them relief. Attending to avoidance commitments when calculating a defendant's culpability allows us to acknowledge at least one important mitigating role of the defendant's environment while maintaining a firm compatibilist commitment to treating people as responsible for their choices, despite their not being able to choose the kind of person they

---

[124] *See* Gary Watson, *Two Faces of Responsibility,* 24 PHILOSOPHICAL TOPICS 227 (1996).
[125] Nagel, *supra* note 56.
[126] *See, e.g.,* Watson, *supra* note 124; T. M. SCANLON, WHAT WE OWE EACH OTHER (2000); Strawson, *supra* note 37; Morse, *supra* note 121; Pamela Hieronymi, *The Force and Fairness of Blame,* 18 PHILOSOPHICAL PERSPECTIVES 115 (2004).

became. But it still leaves open the option of providing a further, separate, diminished responsibility excuse to defendants whose actions, due to environmental circumstances, manifest indifference toward others. For those who are inclined to think that Harris's background should result in mitigation, a mens rea regime focused on avoidance commitments doesn't preclude the option. We would simply need some further diminished responsibility excuse, in addition.

In short, there are at least two fundamentally different ways in which poverty, abuse, and other forms of environmental deprivation can be criminogenic. One causal path is by reducing an agent's concern for others, another is by masking that concern. What attending to avoidance commitments reveals is that we need not agree on what to do with defendants like Robert Harris, in order to provide relief to another uncontroversially less dangerous and less culpable population of defendants suffering from environmental deprivation who *do* care about others, but whose concern is masked by a lack of alternatives.

IV. DOCTRINAL IMPLICATIONS AND AVENUES OF REFORM

So far, I have argued that there is a certain class of purposeful wrongdoer – the reluctant wrongdoer who engages in intentional wrongdoing as a necessary means to some further goal, but who is also committed to attending to and pursuing alternative means toward that goal, even when those alternative means are costly – who should be treated as less liable than particularly callous wrongdoers who commit homicide recklessly or negligently, but who lack such avoidance commitments. I have further argued that many negligent and reckless police-perpetrated homicides – such as the killing of George Floyd by police officer Derek Chauvin – are instances of such particularly callous wrongdoing.

If this argument is correct, then a criminal code like that of Minnesota, where Chauvin was charged, which treats reckless and negligent police-perpetrated homicides (along with all other reckless and negligent homicides) as less liable than the reluctant purposeful homicides of the kind detailed by Browne and Williams, is in need of amendment.

In this Part, I show how the dispositional analysis of culpability I have advanced in terms of avoidance commitments has important doctrinal implications: both for diagnosing how and why current homicide mens rea doctrine fails to properly sort cases of police homicide relative to cases of reluctant purposeful homicide, and for how that doctrine can be refined and revised.

Homicide law is particularly important for criminal law theorists because, due to the severity of the crime, it is the place where mens rea doctrine is developed in the most sophisticated manner. In particular, more than for most other categories of offences in criminal law, criminal homicide regimes are not blind to the fact that the rigid ordinal ranking of intentional harms as more liable than harms of recklessness or negligence may fail to track the underlying culpability and dangerousness of some offenders. Both MPC and common law homicide jurisdictions involve mechanisms both for increasing the liability of certain reckless homicides – those acting with a

"depraved heart" in the common law or "extreme indifference to human life" in the MPC – and for decreasing the liability of certain intentional homicides – intentional homicides committed in the "sudden heat of passion based on adequate provocation" in common law jurisdictions, and intentional homicides committed under "extreme emotional disturbance" in the Model Penal Code.

One might suspect that the problems this Article raises suggest not a radical rethinking of the current mens rea hierarchy, but simply the broader application of these specialized mens rea categories to criminal offenses more generally. However, as I will argue in the rest of this Part, the doctrines of "depraved heart" recklessness and "heat of passion" or "extreme emotional disturbance" fail to capture the features of the respective psychologies of the callous and reluctant agents in virtue of which the former are more culpable and more dangerous than the latter. By focusing on homicide, these cases allow for a discussion of why the current available resources for mitigating and aggravating mens rea carveouts in criminal law are insufficient, and the shape that a more successful doctrine might take.

In subsection A, I will begin by spelling out in more detail the criminal homicide regimes of the common law and the Model Penal Code, including the aggravating and mitigating categories of "depraved heart recklessness" and "provocation-passion/extreme emotional disturbance." In sub-sections B and C, I show that each of these doctrines is inadequate to the task of properly apportioning the respective liabilities of the two classes of homicides, and expand on my diagnosis of why they are insufficient. In Part V, I will then offer some possible prescriptions for improving the current regimes in light of that diagnosis.

IV.A. The Current Homicide Regime(s)

While the Minnesota statutory scheme under which Derek Chauvin was charged is broadly representative of the criminal homicide gradation regime in most states, there are variations across different jurisdictions. Many states' criminal homicide regimes, like Minnesota, follow the Federal Code in using some variant of the common law degree structure based upon the Pennsylvania Reforms of 1794 (the "Penn System") which divides murder and manslaughter into degrees.[127] Of those of states who do not use a variant of the Penn System, most follow some variant of the Model Penal Code, which divides up homicide based upon the Model Code's standardized purpose-knowledge-recklessness-negligence mens rea hierarchy.

In this subsection, I will discuss each system in turn, beginning with the Model Penal Code. I will then work through how neither system can appropriately sort the case of Derek Chauvin and cases of reluctant purposeful homicides like those committed by the domestic violence survivors of Browne & Williams' study.

---

[127] See 2 MODEL PENAL CODE AND COMMENTARIES 16 (AM. L. INST. 1980). As of 2012, Twenty nine states, the District of Columbia, and the federal government continued to follow the Penn System in dividing murder into further degrees based upon premeditation or deliberation. Kimberly Ferzan, *Plotting Premeditation's Demise*, 75 LAW AND CONTEMPORARY PROBLEMS 83, 84 (2012).

i. Homicide in The Model Penal Code and Penn System

Criminal Homicide in the Model Penal Code is shaped by the Model Penal Code's more general division of mens rea elements into four categories of 'purpose,' 'knowledge,' 'recklessness' and 'negligence.'[128] Model Penal Code section 210.1(2) divides criminal homicide into three grades: murder, manslaughter, and negligent homicide. Murder includes ordinary cases of homicides "committed purposefully and knowing;"[129] manslaughter includes ordinary cases of homicides "committed recklessly;"[130] and negligent homicide includes all cases of homicide "committed negligently."[131]

However, in addition to this standard sorting, the Model Penal Code's homicide regime also includes two more specialized mens rea states to provide further, more fine-grained, "culpability requirement in addition to those used more generally throughout the model code."[132] These specialized mens rea states help aggravate certain categories of reckless homicide, and mitigate certain categories of purposeful and knowing homicide, moving them up or down by one grade, respectively.

The first of these specialized mens rea states is "reckless[ness]… manifesting extreme indifference to the value of human life," which the Model Penal Code includes along with knowledge and purpose as sufficient mens rea for murder.[133]

This category can be understood to function as a mechanism to try to capture cases of especially bad reasons-responsiveness that are not appropriately reflected by the division into purpose on the one hand, and recklessness on the other. Indeed, the advisory notes to the Model Code are explicit that this is the category's intended function. In making "the judgment that there is a kind of reckless homicide that cannot be fairly distinguished in grading terms from homicides committed purposefully or knowingly," the Model Penal Code advisory notes make specific reference to the idea that grading should be grounded in degree of insufficient concern.[134] As the advisory note explains, "[t]he significance of purpose or knowledge as a standard of culpability is that, cases of provocation or other mitigation apart, purposeful or knowing homicide demonstrates precisely such indifference to the value of human life."[135] The inclusion of certain reckless cases is due to the Model Penal Code drafters' judgment that in certain cases, "recklessness is so extreme that it demonstrates similar indifference."[136] However, while certain kinds of extreme recklessness can be "assimilated to purpose or knowledge for the purposes of grading," there is no such equivalent aggravating

---

128 MODEL PENAL CODE § 2.02 (AM. L. INST. 1962).
129 MODEL PENAL CODE § 210.2(1)(a) (AM. L. INST. 1962). Excepting those purposeful and knowing homicides committed "under the influence of extreme mental or emotional disturbance for which there is a reasonable explanation or excuse."
130 MODEL PENAL CODE § 210.2(1)(b) (AM. L. INST. 1962). Excepting those cases of reckless homicide committed with "extreme indifference to human life."
131 MODEL PENAL CODE § 210.2(1)(c) (AM. L. INST. 1962).
132 2 MODEL PENAL CODE AND COMMENTARIES 17 Note 37 (AM. L. INST. 1980).
133 MODEL PENAL CODE § 210.2(1)(b) (AM. L. INST. 1962).
134 2 MODEL PENAL CODE AND COMMENTARIES 21 (AM. L. INST. 1980).
135 *Id.*
136 *Id.*

form of negligence in the MPC regime, no matter how severe the negligence is. As the advisory notes emphasize, "the model code provision makes clear that inadvertent risk creation, however extravagant and unjustified, cannot be punished as murder."[137]

The inclusion of a certain kind of recklessness within the highest grade of murder is not, by itself, particularly exceptional. The model Penal Code mens rea hierarchy, as laid out in Model Penal Code Section 2.02, evinces a commitment to a *weak ordering* between purpose, knowledge, recklessness, and negligence, in that a purposeful harm must be treated as *equally or more liable* than a knowing harm, a knowing harm must be treated as *equally or more liable* than a reckless harm, and so on.[138] But it does not require that purpose be treated as *more liable* than knowledge, and so on down the hierarchy. This weak ordering is enshrined in MPC section 2.02(3) in the treatment of ungraded offenses. For ungraded offenses, the Model Penal Code treats recklessness as the default minimum mens rea "unless otherwise provided."[139]

But, as we have already seen, even this system of *weak ordering* is insufficient to properly apportion liability between the non-intentional police perpetrated homicides and the reluctant female perpetrated intimate partner homicides of Browne and Williams' study. The most it could do is treat the callous reckless or negligent agent as *equally liable* to the reluctant purposeful agent. But it cannot accommodate (and is designed to prevent) a grading scheme where certain purposeful agents are *less liable* than certain reckless agents.[140] And the inclusion of "reckless homicide manifesting extreme indifference to human life" in the highest grade of homicide in the Model Penal Code's homicide regime in MPC Section 210 does not change this basic structure. The most it allows is for a subset of reckless agents to be treated as equally liable to the standard purposeful agent. As the advisory notes make clear, section 210 is designed to avoid a regime, like New York's pre-MPC grading system, which "treated some intentional killings less seriously than the class of unintentional killings covered by the 'depraved mind' part of the statute."[141] Indeed, by choosing to provide grades of homicide at all, rather than using the default "recklessness-plus" system of MPC section 2.02(3), the code's treatment of homicide arguably involves treating *fewer* recklessness cases as equivalent to purpose than it would for other generic ungraded offences. So even with the category of 'extreme' recklessness, the Code creates a greater difference in liability between purposeful and reckless murders relative to ungraded offences, not less liability.[142]

---

[137] *Id.*

[138] For a more detailed discussion of the Model Penal Code's mens rea hierarchy, see Antill, *supra* note 27.

[139] MODEL PENAL CODE § 2.02(3) (AM. L. INST. 1962).

[140] As I have previously put the point, the standard weak ordering of the Model Penal Code cannot accommodate cases of *inter-hierarchical* divergence in culpability or dangerousness. Antill *supra* note 27.

[141] 2 MODEL PENAL CODE AND COMMENTARIES (AM. L. INST. 1980).

[142] There is, of course, discretion to further differentiate cases at sentencing. But in most grading regimes, as in the Federal Criminal Code, this discretion is limited, and even when courts do exercise such discretion, choices about departures typically take place in the shadow of baseline sentencing ranges, which are themselves typically set by the mens rea regime. See discussion *supra* note 26 and accompanying text.

Where the Model Penal Code regime appears to provide a more radical departure to the standard mens rea hierarchy is in the inclusion of an additional mitigating factor for certain purposeful or knowing actions committed "under the influence of extreme mental or emotional disturbance for which there is reasonable explanation or excuse."[143] These purposeful and knowing homicides are classified as a form of manslaughter, along with homicides committed with ordinary criminal recklessness which do not exhibit "extreme indifference" to human life.

The Penn system, variants of which are enshrined in the Federal Code in 18 U.S.C. 1111 –1114 and in many states, such as Minnesota, divides homicide even more finely than the Model Penal Code, splitting the grades of murder and manslaughter into further degrees.

Under the Penn system, murder is divided into at least two degrees. The most severe of these is 'first degree murder' which, as in the federal code, is typically defined to pick out a subset of intentional homicides which are "willful, deliberate, malicious, and premeditated killing[s]."[144] Second degree murder requires that the homicide be performed with "malice aforethought"[145] which, as in the Model Penal Code, typically includes both "intentional and knowing homicide."[146] In some variants, the "malice aforethought" requirement also includes homicides performed recklessly while "evincing a depraved mind, without regard for human life." In other variants, like Minnesota, the category of homicides performed recklessly "without regard for human life" constitutes a third, even lessor degree of murder.

Like the Model Penal Code, the Penn system also picks out a subset of purposeful homicides – those performed "in the heat of passion engendered by adequate provocation" to be treated as manslaughter, along with homicides committed with ordinary recklessness and gross negligence. Unlike the Model Penal Code, however, the Penn system divided manslaughter into finer degrees as well. Typically, the system reserved the highest grade of manslaughter – termed first-degree manslaughter or 'voluntary' manslaughter – for mitigated purposeful homicides committed in the heat of passion through provocation. Ordinary reckless and negligent homicides were categorized into lesser grades of involuntary manslaughter, sometimes grouped together, and sometimes, mirroring the Model Penal code, separated into two grades of second- and third- degree manslaughter, respectively.

There are several important differences between these two systems with consequences for the reluctant purposeful agent's liability relative to the liability of callous non-intentional police-perpetrated homicides like Chauvin's killing of George Floyd. Because first-degree murder does not allow for a provocation-passion defense, the inclusion of first-degree murder in the Penn system allows the classificatory system

---

143 MODEL PENAL CODE § 210.(3). (AM. L. INST. 1962).

144 *See* 18 U.S.C. 1111; 2 MODEL PENAL CODE AND COMMENTARIES 17 (AM. L. INST. 1980). With the exception of a few jurisdictions, which include among 'willful and premeditated' homicides knowledge cases where the agent's actions were premeditated and the agent knew to a practical certainty that the action would result in the death of another person.

145 *Id.*

146 *Id* at 14.

to pick out a class of intentional homicides – those that were pre-meditated – which will always involve more liability than the most liable depraved heart recklessness crimes. Second, whereas the MPC system groups purposeful homicides committed under "extreme emotional disturbance" with ordinary reckless homicides, the Penn system still typically assigns even these less culpable purposeful homicides more criminal liability than ordinary reckless homicides (classifying the former as "voluntary manslaughter" and the latter as "involuntary"). Third, while "passion-provocation" and "extreme emotional disturbance" defenses tend pick out similar kinds of cases (those where the emotions or passions overwhelm the agent's rational decision-making), the precise scope of the two doctrines of may differ. The 'provocation-passion' defense of the Penn system was often limited, at least historically, to certain kinds of enumerated provocations, and limited to provocations committed by the deceased toward the defendant. The MPC category is self-consciously broader, expanding the class of permissible "provocations" to include "situations where the provocative circumstance is something other than an injury inflicted by the deceased on the actor but nonetheless is an event that arouses extreme mental or emotional disturbance."[147]

More striking than these differences, however, are the similarities. Both the MPC and Penn system include specialized mens rea states for increasing liability for reckless homicides which seem to manifest particularly large failures to respond to the value of other people, and both include specialized mens rea states for decreasing liability for a subclass of purposeful homicides involving extreme passion or emotion. Combined, these two mens rea states appear to allow what is otherwise absent from the criminal code: the elevation of certain cases of reckless wrongdoing (reckless homicide exhibiting extreme indifference to human life) *above* certain cases of purposeful wrongdoing of the same kind (purposeful homicide under extreme emotional disturbance). Here, a subset of reckless crime appears not just equally, but *more* liable, than a subset of purposeful crimes.

It may appear that this kind of carve-out is precisely the kind we would want to capture both the reluctant purposeful homicides of domestic violence survivors and the callous reckless police-perpetrated homicides of officers like Derek Chauvin. But, as we will see in the following sub-sections, the current doctrine is not actually particularly well suited to properly sorting either of these cases of homicide. The mitigating defenses of both the MPC and the Penn system are designed to accommodate different exculpatory features than the features of the reluctant purposeful agent. In particular, they are designed to capture cases where the agent's action does not 'really' reflect their actual choices, by picking out cases where the agent's rational will is 'overwhelmed' by an emotion, so that they are alienated from the resulting action, and so not 'really' intending after all. Because of this focus on the dichotomy between 'reason' and 'passion' or 'emotion', this defense is unable (I will argue below) to provide effective mitigation of criminal liability for the reluctant, but reasoned and often premeditated, acts of intimate partner homicide by survivor-defendants trapped in abusive relationships. In fact, as many critics have noted, the 'emotion/passion' defense is instead far more likely to give reduced liability to *abusers*

---

[147] *Id.* at 49.

who murder their victims, than to survivors of that abuse, since "adultery" has long been one of the categories of "passion-provocations" recognized by the courts, and has for just as long been used historically by the courts as an excuse to provide leniency to violence perpetrated against women by abusive spouses.[148]

ii. How Current Doctrine Mis-Sorts Homicide Liability

In the prior subsection I have described the two dominant homicide regimes in the United States. The hope of this Article is that, by attending to the intentional commitments of purposeful, reckless, and negligent agents which are particularly exculpatory or particularly problematic, we can make headway in evaluating the respective merits and shortcomings of each of these systems.

Before moving on to diagnosing *why* the current system is inadequate, I will motivate the need for such diagnosing by showing in this subsection how neither of the two current regimes is capable of correctly apportioning liability to Chauvin relative to survivor-defendants. By working through in detail how the current regime applies (or misapplies) to these cases, we can bring to light some of the deeper problems with the current regimes' approach more generally.

Consider first the case of Derek Chauvin's act of homicide in killing George Floyd. On the analysis of callous agents in terms of avoidance commitments that I have offered in Parts I-III, the question of whether Chauvin was aware of a risk that his actions might cause George Floyds death, or whether he was unaware of such a risk, is irrelevant to the real source of his extra culpability. What makes such callous agents both so culpable and so dangerous is a special lack of concern for the victim, constituted, at least in part, by the absence of any commitment to seeking out or pursue alternate means to their goals (for Chauvin, effecting Floyd's arrest) which avoid harm to others. The absence of any such commitment at least partly constitutes the sense in which Chauvin failed to treat Floyd's life as though it mattered. And that absence of a commitment to avoid harm is often manifested as clearly, if not more clearly, by negligence than it is by recklessness.

This difference in subjective awareness, however, makes a great deal of legal difference in the current doctrinal system. If it is true that Chauvin was subjectively aware of a risk that his actions might cause Floyd's death, Chauvin was criminally reckless. Whatever that subjective probability of death was, disregarding it to affect an arrest more easily would involve "a gross deviation from the standard of conduct that a law-abiding person would observe in the actor's situation," the standard definition of recklessness from the Model Penal Code.[149] If Chauvin was unaware of the subjective risk, he would instead be criminally negligent: Chauvin's "failure to perceive it…involves a gross deviation from the standard of care that a reasonable person

---

[148] *See, e.g.* Victoria Nourse, *Passion's Progress: Modern Law Reform and the Provocation Defense*, 106 YALE L.J. 1331, 1332 (1997); Rozelle, *supra* note 63 at 197; Laurie Taylor, *Provoked Reason in Men and Women: Heat-of-Passion Manslaughter and Imperfect Self-defense,* 33 UCLA L. REV 1679, 1679 (1986).

[149] MODEL PENAL CODE § 2.02(2)(c) (AM. L. INST. 1962).

would observe in the actor's situation," the standard definition of negligence from the Model Penal Code.[150]

Because neither the Penn System nor the Model Penal Code allow for the possibility of aggravated liability for depraved heart negligence, if Chauvin was negligent, he would be liable only for involuntary manslaughter (under the Penn system) or negligent homicide (under the MPC).

Supposing Chauvin was reckless, he might be subject to increased liability if his recklessness counts as aggravated recklessness under the Penn "depraved heart recklessness" doctrine or the MPC "recklessness manifesting extreme indifference to human life" doctrine.

Since the MPC regime treats the question of whether recklessness manifests extreme indifference to human life as a basic concept that "can[not] be further clarified" and "must be left directly to the trier of fact,"[151] there is no barrier to holding Chauvin liable for extreme recklessness. Chauvin would thus be liable for murder under the MPC.

For many regimes following the Penn system, in contrast, the difference between 'ordinary' recklessness and 'depraved heart' recklessness is typically a matter of the degree of probability of the victim's death (such as whether the probability was higher than 50%).[152] This is sometimes cached out in terms of the subjective probability assigned to the death, and sometimes cached out in terms of objective probability of the death.[153] On such regimes, it is still not clear that Chauvin's act would properly qualify for aggravation under the 'depraved heart recklessness' doctrine. Chauvin may well have thought the subjective probability of Floyd's death was quite small (if he assigned it any possibility at all). What makes his act so culpable – and the sense in which he failed to give sufficient concern to Floyd's life – was in failing to give that probability (small or large) its due weight.

Even if Chauvin were found to satisfy the requirements for depraved heart recklessness on the Penn system, he would not be guilty of first-degree murder. Because the murder was not intentional, it could not be premeditated. Chauvin would be liable for either second-degree murder (as in the Federal system) or third-degree murder (as in the Minnesota system) depending on which variant of the Penn system is enshrined, and where cases of 'depraved heart recklessness' are placed.

Consider now the liability for cases of female perpetrated intimate partner homicides under contemporary homicide regimes, or other reluctant purposeful homicides committed under exigent circumstances. Because the homicides are committed purposefully, they would be categorized as murder under both the Penn

---

150 MODEL PENAL CODE § 2.02(2)(d) (AM. L. INST. 1962).
151 2 MODEL PENAL CODE AND COMMENTARIES 22 (AM. L. INST 1980).
152 See, e.g., INDIAN PENAL CODE S299 (1860) (requiring for DHR that the actor have "knowledge that he is *likely* by such act to cause death) (emphasis added).
153 MODEL PENAL CODE commentary for § 210.(3). (AM. L. INST. 1962).

and MPC system, absent a finding of "extreme passion" or "extreme mental or emotional disturbance."

But the facts of reluctant purposeful homicide cases cannot be fit easily inside the lines of those two doctrines. Firstly, exigent circumstances, of the kind that might be severe enough to lead an agent in those circumstances to kill another person, when such a killing is a necessary means to escape those circumstances, need not inflame the passions in order to demonstrate reduced ill-will on the part of the offender. Whether the reasons are strong enough to inflame the passions is, at root, a psychological inquiry. But moderately strong reasons for acting are consistent with moderately strong concern for their victim, and there is no *a priori* reason why those moderately strong reasons would have to be the kind of thing that excite the passions or disturb the emotions.

Secondly, the degree of ill will exhibited by reluctant purposeful actors doesn't change depending on whether the action was done immediately, or after the kind of 'cooling off period' which negates the typical extreme emotional disturbance or passion-provocation defense.[154] The domestic abuse survivor who commits intimate partner homicide in Browne and Williams study, for example, might both have planned out the homicide for some time, and also have been committed to abandoning the plan and taking much more costlier alternatives to that homicide, like leaving her home to go to a shelter, if such shelters had been available.

The fact that such reluctant purposeful homicides were the result of a pre-meditated and deliberate choice is still consistent with a commitment to take alternatives when they open up, and so consistent with (relatively) little malice or insufficient concern toward the victim on the part of the defendant. It is thus the kind of purposeful homicide which merits some partial mitigation. But the fact that such homicides are, or could be, premeditated, ensures that the reluctant purposeful homicides will still be classified as murder on both the MPC and the Penn system, rather than downgraded due to extreme emotional disturbance or provocation.

Thus, even with "extreme emotional disturbance" and "provocation-passion" as mitigating excuses, and "depraved heart recklessness" and "extreme indifference" as aggravating factors, current doctrine lacks the resources to make police perpetrated homicides like Chauvin's killing of George Floyd more liable than purposeful homicides like those in Brown & Williams' study. Under the Model Penal Code, at best, a reckless police homicide will be treated as *equally* liable to the reluctant purposeful homicide (both will be classified as murder). At worst, a negligent police homicide will be treated as substantially less liable (the police-perpetrated homicide will be classified as negligent homicide, the reluctant purposeful homicide as murder). Under the more common Penn system, things are even worse. The reluctant purposeful homicide will be liable for first-degree murder. In contrast, the police perpetrated homicide will be liable for, at worst, second- or third- degree murder

[154] See, e.g., People v. Shelton, 385 N.Y.S. 2d 708 (N.Y. Sup. Ct. 1976), whose definition of extreme emotional disturbance has been adopted by most states that use EED rather than provocation-passion, which emphasizes the importance of the time interval in determining whether EED applies.

(depending on how depraved heart recklessness is categorized) and, if negligent, will be liable only for involuntary homicide.

IV.B. The Inadequacies of the Extreme Emotional Disturbance/Passion-Provocation Doctrines

The central problem with the provocation-passion and extreme emotional disturbance doctrine, as well as more recent affirmative defenses like "battered persons syndrome" or "coercive control" is that all such doctrines still implicitly treat intentional wrongdoing as the paradigm expression of ill-will. Cases of "heat of passion" or "extreme emotional disturbance" are understood as cases where the action is not *fully* intentional. The mitigating excuse is that, due to the extreme emotional disturbance or heightened passion, the agent is acting in ways that their autonomous selves would not endorse, nor intend, had they been in control of themselves and properly self-governing.[155]

Underlying the current provocation-passion/extreme emotional disturbance doctrine is a particular, and outdated, view of our moral psychology. The idea is that an agent who is not in the grip of passion, or extreme emotional disturbance, controls their decisions by exercising their reason to deliberate, weighing the reasons for and against acting. When an action is a product of such deliberate reasoning, the action reflects the agent's take on how important those reasons are. It is therefore expressive of their will.

In contrast, when an agent is in the grip of a passion or extreme emotion, while they may still act intentionally[156] – that is, the action might be conscious and aim-directed – it will be an intention that reflects their impulses, rather than their 'true' will or 'true' attitudes and values.[157] It is the "emotion" doing the acting, so to speak, rather than the agent themselves. It is for this reason that "provocation" is typically classified with "diminished responsibility" defenses.[158]

The problem with such "diminished responsibility" excuses, as shown by attending to the importance of avoidance commitments, is that they are the wrong kind of excuse to capture the mitigating features of the reluctant agent. Following Peter Westen's influential framework, we can understand excuse as functioning to show that

---

[155] This view is perhaps most clearly articulated in Stephen Garvey's account of provocation as *akrasia*, or weakness of will. Stephen P. Garvey, *Passion's Puzzle*, 90 IOWA L. REV. 1677 (2005).

[156] *See* MPC commentary Vol 2 at 54 ("A sudden rage, however engendered, does not necessarily or even probably negate an intent to kill. More likely it reinforces the firmness of the actor's resolve to take the life of another. At most, therefore, provocation affects the quality of the actor's state of mind as an indicator of moral blame- worthiness. Provocation is thus properly regarded as a recognition by the law that inquiry into the reasons for the actor's formulation of an intent to kill will sometimes reveal factors that should have significance in grading.").

[157] See SUSAN WOLFE, FREEDOM WITHIN REASON (1993) for the seminal discussion of such "deep self" or "real self" views in philosophy.

[158] *See also* H.L.A. Hart, *Negligence, Mens Rea, and Criminal Responsibility in* PUNISHMENT AND RESPONSIBILITY 153 (2008) ("In these last cases, exemplified in 'provocation' and 'diminished responsibility', if we punish at all we punish *less,* on the footing that, though the accused's capacity for self-control was not absent its exercise was a matter of abnormal *difficulty.* He is punished in effect for a failure to exercise control.").

the defendant is not actually an "apt object of" negative reactive attitudes, such as reproach, by showing that the defendant did not actually "act[] with a reprehensible attitude toward the legitimate interests of... others."[159]

One way to show that a defendant is not an apt target of reproach is to show that the action in question does not really reflect the agent's attitudes at all. This is clearest in failures of the voluntary act requirement: if the agent is, e.g., sleep walking, when they pushed the victim down the stairs, an action which might *seem* to reflect ill-will toward the victim, in fact reflects nothing of the defendant's will at all, because it was not actually intended *by the agent*.[160] Though "partial responsibility" excuses do not deny that the action was voluntary in the same was as a somnambulism excuse, they are making a similar style of claim about the absence of connection between the agent's attitudes and their action (it was the agent's emotions, or mental illness, that did the acting, so to speak, not the agent themselves).

This is, as we have just seen, the way theorists and legislators have understood the passion-provocation and EED excuses. If the action was not fully under the agent's voluntary control – if it was the result of passion, not reason, and so not reflective of a genuine choice – then it is not really reflective of their judging the triggering provocation as more valuable than the well-being of the victim. The fact that they were not in control of themselves shows that the action does not really manifest the ill will which such actions would typically manifest, by showing that the action is not (directly) expressive of the agent's will at all.[161] As the MPC notes, this also explains why provocation/passion is not available for "deliberate and premeditated" intentional homicides which are categorized as first degree murder. If they are a product of deliberation, they are necessarily fully reflective of the agent's will. It similarly explains why such defenses nearly always require the absence of a "cooling off" period. If the agent had a chance to cool off, then their subsequent action reflects their reasoned view of the respective value of their victim and their goal, rather than merely being the product of an intervening emotion.

Finally, the focus on "diminished responsibility" when mitigating purposeful offenses explains why even more recent defenses that are specifically tailored to the case of domestic abuse survivors, such as "battered persons syndrome" often still fail to capture those cases most in need of mitigating. Acknowledging the inadequacy of extreme emotional disturbance/provocation passion doctrines to capture the appropriate reduction of culpability in such cases, some jurisdictions have attempted

---

159 Peter Westen, *An Attitudinal Theory of Excuse,* 25 LAW & PHIL 289 (2006).

160 *See, e.g., R v Parks*, [1992] 2 S.C.R. 871.

161 One might ask: if they are not volitional, why treat them as lesser crimes, rather than not as crimes at all? The MPC's answer seems to be that insofar as the purposeful agent in the heat of passion is culpable, then, they are culpable for a very different kind of failing than the standard intentional agent. They are culpable for failing to 'reign in' their passions the way an agent who cared more about the victim would have (along with a culpable "extraordinary susceptibility to intense passion" and "extraordinary weakness of reason and consequent ability to bring such desires into play"). A full analysis of the soundness of the MPC's moral psychology is the task of another paper. It suffices for this project that, whatever mitigating consequences for culpability the MPC thinks are exhibited by such deficits, rather than the deficits of a standard intentional actor, they are not the deficits of the reluctant purposeful actor.

to provide specialized affirmative defenses, such as defenses of "battered persons syndrome" or "coercive control" for cases of homicides committed by survivors in certain cases of extreme abuse by the victim. Critics have rightly identified ways that using "battered persons syndrome" as a legal model for decreasing liability involves a problematic way of framing survivors' experiences as pathological.[162] We can now see that the basic assumptions of the traditional mens rea hierarchy can help explain some of the legal pressure for such problematic models: the focus on intentional or deliberate wrongdoing forces defendants to either deny their agency through defenses like battered person syndrome, or else face maximal criminal liability for first degree murder. Whereas a system of excuse that recognizes that the deliberate purposeful commission of a crime can nonetheless involve relatively little culpability could allow defendants a way of acknowledging both their agency and their reduced culpability, by accurately explaining how the two are compatible.

And, importantly, such a system of excuse is at least theoretically available. Diminished responsibility is only one of at least two ways that an intentional action may exhibit less ill-will, or malice, than it typically would. In addition to showing an absence of ill will by showing the action doesn't reflect one's will at all, one can also show that the action reflects one's will, but that the will isn't as ill as it might initially appear.

The second kind of excuse shows an absence of ill will by accepting that the action is reflective of the agent's will, but that the agent's will exhibits more concern for the victim than it initially appeared to. Though less often self-consciously identified as such, these excuses are already present in the law. If the defendant, suffering from a mistake of fact, for example, falsely believes that the gun they are firing is empty, their action – shooting the victim – also does not exhibit the ill will of a typical murder. Such agents are excused, not because their action doesn't reflect their will (they were fully deliberate in deciding to pull the trigger) but because they did not really undervalue the life of their victim the way it appeared they did, when it appeared they believed their actions would result in the victim's death.

Understanding the mitigating aspect of reluctant purposeful agents in terms of avoidance commitments helps make it clear that what is needed to capture the reluctant purposeful actor is this second kind of excuse. Their actions reflect less ill will, not because the action does not reflect *their* will, but because their will, while fully manifested in the action, is actually less malicious than the will of most non-reluctant purposeful actors. The commitment to pursuing costly alternatives when available, and the strength of their countervailing reasons to act, show that the reluctant purposeful agent is manifesting a will that involves a relatively higher degree of concern for others.

But by requiring defendants to show that their actions were the result of diminished responsibility, we make such an excuse legally inaccessible. The current doctrine will thus be both under- and over-inclusive. It is under-inclusive in the sense

---

[162] See Cheryl A. Terrance, Karyn M. Plumm, & and Katlin J. Rhyner, *Expert Testimony in Cases Involving Battered Women Who Kill: Going Beyond the Battered Woman Syndrome*, 88 North Dakota Law Review 921, 940-952 (2012), for a recent historical survey of this critical literature.

that it will make it impossible for agents whose actions are fully intended, but who would have sought alternative means, to make this excuse, instead either precluding them (via 'premeditation' or 'cooling-off' exclusions) from offering the excuse at all, or forcing them to claim instead that they were suffering from mental illness or emotion, instead of acknowledging that their action was truly the deliberate result of their will, but that their will (while culpable) was less culpable than that of other homicides. And it will be over inclusive in that it *will* excuse many purposeful agents (like the abusive spouse who murders his partner in a rage) for whom homicide was *not* a necessary means, by arguing that their actions did not reflect their 'true' values.

IV.C The Inadequacies of Depraved Heart/Recklessness with Extreme Indifference Doctrines

In contrast to the mitigating categories of extreme-emotional disturbance and passion-provocation, the doctrines of Depraved Heart Recklessness/Recklessness with Extreme Indifference to Human life (DHR/REI) are closer to successfully capturing the underlying grounds for increased liability. The defense acknowledges that there is something especially culpable about certain mens rea states that fall short of purpose, and that what makes them especially culpable is that they can show the same ill-will as typical purposeful actions.

There are, however, two major problems with the doctrines as they stand. The first problem is that the current DHR/REI doctrines fail to go far enough, along both a vertical and horizontal dimension. The doctrine is too narrow 'horizontally' in the sense that not all cases of callous homicide which manifest an absence of avoidance commitments will qualify under DHR/REI doctrine. By restricting cases of 'extreme indifference to human life' to cases of recklessness, but excluding cases of negligence, the current doctrine fails to recognize that negligence, just as much as recklessness, can manifest the kind of extreme indifference to human life that is characterized by a lack of commitment to seeking out and taking alternative means that do not lead to the victim's death. Vertically, the doctrine is too narrow in the sense that even for those especially culpable callous defendants who *are* captured by the current legal regime, DHR/REI fails to make enough difference in criminal liability to ensure that liability is properly apportioned according to culpability. While these categories can 'bump up' liability by one grade, and so treat cases of extreme recklessness as (at their strongest) *equivalent* to certain cases of purposeful homicide, it will never treat them as *more* liable than cases of ordinary purposeful homicide, and so fail to capture the ways in which homicides like that of Chauvin, who value their victims' lives so little that they are entirely unmotivated to seek out less harmful alternatives, *are* more culpable. As we have seen, the Model Penal Code was designed self-consciously to ensure this result.[163] Even worse, in the Penn system, DHR/REI typically fails to be treated as even *equally liable* to deliberate or premeditated intentional homicides. It is categorized, instead, as second-degree or third-degree murder, with first-degree murder reserved for cases of intentional premeditation. But since, as we saw in the previous section, many of the (relatively) less culpable purposeful homicides may well be deliberate or premeditated, this means that DHR/REI defenses, as they currently stand, will not go nearly far enough to eradicate the misapportionment of liability.

---

[163] *See supra* note 141 and accompanying text.

The second and more serious problem is that the current DHR/REI doctrine, while acknowledging that there is a class of reckless homicides which are especially culpable, lacks a way of appropriately identifying them. Thus regimes tend to either classify DHR/REI cases using what we can now see is some irrelevant measure, such as whether the defendant assigned a degree of probability greater than 50% to their action causing the death of their victim,[164] or, as with the Model Penal Code, leaving the question to the discretion of the fact-finder without any direction about what specific psychological features would make one reckless agent more culpable than another.[165] The problem with this approach, as with Justice Potter's "I know it when I see it" approach to obscenity, is that it would allow (in fact, demand) enormous discretion on the part of fact-finders to make case-by-case decisions about which defendants seem particularly bad to them, when they are ill-suited to such tasks.[166] As has been well documented, this approach has a number of deeply problematic shortcomings. At worst, it raises serious issues of potential racial- and gender- bias.[167] At best, in will lead to different treatment of similarly situated offenders as different juries employ different moral worldviews about which motives are or are not especially blameworthy.[168] Even putting aside questions of accuracy, such a system will inevitably require factfinders to make decisions about which psychological states to criminalize to what degree which ought to be the proper province of the State, not the fact-finder.[169]

V. INTENTIONAL COMMITMENTS, MOTIVES, AND MENS REA REFORM

In Parts I-III, I have shown that the traditional PKRN mens rea regime of the Model Penal Code and contemporary criminal law threatens to systematically assign disproportionate liability to reluctant purposeful agents relative to callous reckless or even negligent agents. It does so because it fails to recognize the existence and importance of 'avoidance commitments' which are manifested in the actions of reluctant purposeful agents, but which often fail to be manifested in the actions of particularly callous reckless or negligent agents. Moreover, in Part IV, I have shown

---

[164] *See supra* note 152 and accompanying text.

[165] 2 MODEL PENAL CODE AND COMMENTARIES 22 (1980) (AM. L. INST). Rather than define extreme indifference to human life in more specific terms, the Code instead claims that the question of "whether a reckless agent indifference is not a question, it is submitted, that can be further clarified" and so consequently "must be left directly to the trier of fact under instructions which make it clear that recklessness that can fairly be assimilated to purpose or knowledge should be treated as murder and that less extreme recklessness should be punished as manslaughter."

[166] See Mark D. Alicke, *Blaming Badly*, 8 J. COGNITION & CULTURE 179 (2008); Mark Alicke, *Culpable Control and the Psychology of Blame*, 126 PSYCHOLOGICAL BULLETIN 556 (2000); Janice Nadler & Mary-Hunter McDonnell, *Moral Character, Motive, and the Psychology of Blame*, 97 CORNELL L. REV. 255, 270 (2012); Janice Nadler, *Blaming as a Social Process: The Influence of Character and Moral Emotion on Blame*, 74 LAW AND CONTEMPORARY PROBLEMS 1 (2012).

[167] *See, e.g. See e.g.* Chris Guthrie, et al., *Blinking on the Bench: How Judges Decide Cases*, 93 CORNELL LAW REVIEW 1 (2007*);* Jeffrey J. Rachlinski, et al*., Does Unconscious Racial Bias Affect Trial Judges?* 84 NOTRE DAME L. REV. 1195 (2009); Brian Nosek & Rachel G. Riskind, *Policy Implications of Social Cognition* 6 SOCIAL ISSUES AND POLICY REV. 112 (2011)*;* Samuel R. Sommers*, On Racial Diversity and Group Decision Making* 90 JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY 597 (2006).

[168] *Id.*

[169] *See, e.g.,* DOUGLAS HUSAK, OVERCRIMINALIZATION: THE LIMITS OF CRIMINAL LAW (2007).

how the traditional mens rea regime's treatment of intentional wrongdoing as paradigm cases of culpability ensures that current homicide doctrine is unable to properly apportion liability to reluctant purposeful agents relative to callous reckless and negligent agents, even when it acknowledges that the culpability of defendants might diverge from the standard PKRN mens rea categories. Its reliance on a mistaken normative picture of culpability means that current doctrine misidentifies both the *degree* to which that divergence can occur, and the *frequency* with which the divergence will occur.

In the case of purpose, agents can exhibit relatively minor culpability, not only because they are only partially responsible for their actions (e.g., when their intentions are the product of passion, rather than reason), but because their actions, even when deliberate, pre-meditated and fully reflective of their will, reflect a will that isn't actually especially culpable (as in the case of reluctant agents who engage in purposeful wrongdoing only when it is a *necessary* means to their ends, and are committed to avoiding wrongdoing when not necessary). Whereas in the case of non-purposeful agents, callous agents, even when negligent, can exhibit a degree of insufficient concern that can rival, or even exceed, that of a typical premeditated wrongdoing.

How, in light of these problems, might the traditional PKRN hierarchy of the Model Penal Code, and attendant mitigating/aggravating mens rea categories, be reformed? One possible solution, variants of which I and others have suggested in the past,[170] is to simply forgo grading by the traditional mens rea categories, with specialized aggravating and mitigating carve-outs, and instead attempt to assess the agent's malice, degree of concern, or quality of will, directly. In the case of homicide, this could be achieved by removing grading altogether. If we had a general crime of homicide, requiring a minimum mens rea of negligence, with a wide sentencing range, judges could make determinations at sentencing about the degree of malice involved in the respective motives of the different agents, and assign punishment accordingly, regardless of where the agent falls in the traditional PKRN hierarchy.[171]

The central difficulty with such a solution, however, is that absent further guidance, it would simply compound the problems with the current discretion involved in the MPC 'depraved heart recklessness' category, which forces fact-finders (whether the judge or the jury) to engage in case-by-case normative determinations of which motives are particularly morally blameworthy.[172] Such a solution would move

---

[170] *See, e.g., See, e.g.,* Michael Serota*, Guilty Minds,* 82 MD. L. REV. 670 (forthcoming); Kimberly Ferzan, *supra* note 127; ALEXANDER AND FERZAN *supra* note 35; Antill *supra* note 27; WILLIAM J. STUNTZ, THE COLLAPSE OF AMERICAN CRIMINAL JUSTICE (2011); SAMUEL H. PILLSBURY, JUDGING EVIL: RETHINKING THE LAW OF MURDER AND MANSLAUGHTER (1998).

[171] Kimberly Ferzan has proposed something very close to this picture (though with a floor of recklessness, for reasons discussed *supra* note 90 and accompanying text) in Ferzan, *supra* note 127 at 103 ("many jurisdictions currently punish reckless homicides that manifest extreme indifference to human life as second-degree murder. It is this very standard, however, that ought to be used to assess the worst killings, for purposeful crimes of identity and crimes of indifference can both manifest extreme indifference. Although courts will no doubt struggle to give some hard edges to this standard (indeed, they already do), at least they will be working with the right standard.").

[172] In the case of a category like depraved heart recklessness, the malice determination falls to the jury rather than the sentencing judge. However, there is a growing body of empirical evidence that both

us back toward the 'general intent' mens rea standards common to older pre-MPC mens rea regimes.[173] The problem with such regimes is that they created enormous injustices, as fact-finders idiosyncratic biases could lead to wildly varying sentences for similarly situated defendants, depending on what motives those fact-finders happened to particularly dislike. Indeed, the central advance of the Model Penal Code is typically held to be that it directs fact-finders to engage in factual investigations about the presence or absence of certain well-defined psychological states, specified by a legislature with the democratic legitimacy to make such judgments, as deserving of more or less criminal liability.[174]

Instead of having judges or juries make case-by-case gestalt determinations about which agents' minds or motives are particularly malicious, what we would want, ideally, is for the legislature to pick out certain psychological states, described in normatively neutral terms, which the fact-finder could then identify in a factual inquiry into the defendant's psychology. Such a system would be both more democratically legitimate and more likely to result in accurate culpability determinations.

The challenge, of course, is finding the right psychological states. The PKRN mens rea regime, which directs fact-finders to identify if defendants had certain beliefs and intentions, succeeds in defining normatively neutral psychological states, but picks out states which only roughly track the agents' actual subjective culpability. Whereas a psychological state like 'malice,' a 'depraved heart,' or a 'wicked motive,' picks out a category which is by definition more culpable, but is too ill-defined to avoid forcing the fact-finder to make moralized judgements instead of purely factual ones.[175]

We appear faced, in other words, with a variant of the well-known rules-standards dilemma.[176] Criminal law appears to face an unavoidable tradeoff. On the one hand, it can employ a mens rea regime articulated in terms of the standard PKRN mental states, which allow the law to formulate (relatively) straightforward legal rules,

---

judges and juries are susceptible to bias when making such normative determinations. *See* Rachlinski et al. *supra* note 167.

[173] *See, e.g.,* Paul H. Robenson & Jane A. Grall, *Element Analysis in Defining Criminal Liability: The Model Penal Code and Beyond, 35* STANFORD L. REV. 681 (1983).

[174] *Id. See also* Kenneth W. Simons, *Should the Model Penal Code's Mens Rea Provisions Be Amended?*, 1 OHIO ST. J. CRIM. L. 179, 180 (2003) (referring to the MPC element analysis approach to mens rea as a "tremendous advance" over the older common law approach.

[175] Other scholars have advocated an expansion of those regimes which attempt to enumerate some limited number of motives – pecuniary interest, or hatred of a particular racial or ethnic group – as particularly culpable, and then directed the fact-finder to make a factual determination of whether those motives were present. *See, e.g.,* SAMUEL H. PILLSBURY, JUDGING EVIL: RETHINKING THE LAW OF MURDER AND MANSLAUGHTER (1998). At least one distinctive difficulty with such an approach is that if insufficient concern consists in a failure to weigh the value of the victim against the agent's motive (whatever that motive is), then any number of motives could manifest extreme insufficient concern. There is no way to capture *all* cases of insufficient through the enumeration of particular classes of motives without providing fact-finders with precisely the kind of wide-ranging discretion to judge motive which we were trying to avoid.

[176] A problem acknowledged by advocates of discretionary mens rea standards. *See* ALEXANDER AND FERZAN, *supra* note 35, at 264. I believe this framing is useful though I suspect the core issue is less about whether a legal principle is articulated in terms of rules or standards, and more about whether the standard involves a purely factual inquiry, or requires moral discretion on the part of the fact-finder.

at the cost of mis-sorting culpability when those states inevitably fail to comport with the agent's underlying quality of will. On the other hand, the law could employ more flexible discretionary legal standards in terms of morally-laden concepts like 'malice' or 'ill-will,' at the cost of bias and inconsistency across similarly situated defendants.

Criminal law theorists tend to line up on one side or the other of this dilemma, depending on how severe they believe the gap is between the current hierarchy and the actual culpability of defendants. On the one hand, those who believe, like Ken Simons, that the current mens rea hierarchy "works fairly well in translating underlying normative approaches [to] blameworthiness . . . into doctrinal requirements" tend to be attracted to the more rigid, rule-like formulations.[177] Whereas those like Larry Alexander and Kimberly Ferzan, who are more concerned that the gap between the current regime and the underlying culpability of defendants is greater, hold that "in almost all cases the criminal law should opt for standards."[178]

In this final Part of the Article, I want to show how the proposed dispositional analysis of culpability in terms of the presence or absence of avoidance commitments has the promise to sidestep this dilemma altogether, and breathe new life into this intractable debate. The Article's analysis not only has the advantage of giving us theoretical clarity about the degree (and source) of underlying culpability of the indifferent and reluctant agents, it also, in doing so, gives us the resources to craft a normatively neutral, relatively clear standard for fact-finders to apply in lieu of, or at least as a supplement to, the standard PKRN hierarchy.

The purpose of this Part is merely to illustrate in general terms the advantages of a code that attends to avoidance commitments, not to craft a code itself. But a brief look at how such avoidance commitments might be incorporated into the existing homicide regime can help illustrate both the method's appeal and its simplicity. Attending to the agent's intentional commitments provides us with a system of specifying a finely tailored excuse that will allow us to pick out, with more selectivity and sensitivity, those defendants among a PKRN class who are particularly more (or less) culpable.

In many ways, such a system could look quite similar in structure to current homicide regimes. The criminal code could keep its current grading system, but replace (or augment) the current mitigating categories of "purposeful homicide under extreme emotional disturbance" and "purposeful homicide in the heat of passion" with the mitigating category of "purposeful homicide manifesting reluctance." defined in terms of the presence, or absence, of a commitment to seek out and engage in alternative means when available (and where even defendants who engaged in pre-meditated or deliberate intentional wrongdoing could qualify for mitigation). To determine whether a defendant qualified, fact-finders could either consider evidence that a defendant actually sought out and tried to take alternatives before engaging in crime, or (in the absence of any actually available alternatives), evidence that a defendant counter-factually would have tried to take such alternatives if they had been available.

---

[177] Simons, *supra* note 42 at 490.

[178] ALEXANDER AND FERZAN, *supra* note 35, at 264.

In the other direction, we could amend or replace the current "recklessness manifesting extreme indifference to human life" doctrine with a new aggravating category for any homicide that manifested an absence of avoidance commitments. Rather than leave the question of when a negligent or reckless act manifests extreme indifference to human life as a "question to be left to the fact finder," such a code would instead direct the fact-finder to determine whether the reckless or negligent agent failed to exhibit avoidance commitments, by determining whether the means they chose to achieve their goal, which they knew or should have known had a risk of causing the death of another person, was a necessary one. If it could be established beyond a reasonable doubt that the agent either actually ignored, or was counterfactually disposed to ignore, alternative means, juries would be directed to categorize them as a case of homicide manifesting extreme indifference to human life, which could move them up to the most severe grade of murder in the jurisdiction. Because this absence of avoidance commitments can be manifested by negligent, reckless, knowing, or purposeful agents, it could capture the genuinely culpable and dangerous purposeful wrongdoers (those who are not reluctant), as well as the cases of truly culpable and dangerous callous negligent agents.

This is, of course, just one way to implement such a system, and even then only sketched in broad strokes. Rather than codify avoidance commitments in an affirmative defense or a mitigating carve-out to the general PKRN hierarchy, a more ambitious system might eschew the PKRN hierarchy entirely. This stronger version of the proposed reform is not that we should invert the PKRN hierarchy, and so treat negligence and recklessness as more liable than purpose. Rather, it is that we should replace the current hierarchy entirely, and focus instead on intentional commitments, as evidenced by a defendant's actual and counterfactual dispositions to avoid wrongdoing.

In such a system, we could get rid of the four familiar mens rea states of purpose, knowledge, reckless, and negligence, and replace it instead with a general mens rea regime which calculated culpability in terms of counterfactual dispositions. This could be a bipartite system, with one counter-factual test, to determine whether a defendant manifested avoidance commitments and sort their degree of culpability accordingly. Or it could be a more complicated system using several counterfactual tests, capturing a variety of important overlapping intentional commitments, to divide mens rea into three or four categories, with a counterfactual test for avoidance commitments playing the central role that purpose plays in our current system of criminalization.[179]

---

[179] Spelling out such a system in detail is beyond the scope of this paper (though I hope to pursue the topic in future work). I think (for the reasons outlined in this paper) that avoidance commitments would be central for such a system. Another possible auxiliary intentional commitment could be the civil law counterfactual test for *dolus eventualis*, used in many European jurisdictions such as Germany and Italy. A mens rea system using counterfactual tests could even give some auxiliary role to purpose by including a counterfactual test for the presence or absence of tracking commitments (which, as we have seen, are constitutive of purpose).

Since, as I have argued in Part III, we should expect proportionately more negligent and reckless agents to be callous, and proportionately more purposeful agents to be reluctant, this will have the practical effect of subjecting many purposeful agents to less liability than their reckless counterparts. Such a system would allow us to group together any callous purposeful agents who lack avoidance commitments (like the purposeful assassin) with callous reckless and negligent agents (like Chauvin) who lack avoidance commitments. But it would also allow those reckless agents who *do* express avoidance commitments (like the reluctant drug dealer who foresees but does not intend a risk of death to his clients) to receive appropriate mitigation as well.

It would also provide an avenue for recognizing the exculpatory reluctance of many socially disadvantaged defendants who commit crime only as a last resort, without requiring that we treat social disadvantage as an independent mitigating factor. This would allow the system to provide relief from excessive criminal liability to the uncontroversially less culpable and less dangerous population of reluctant purposeful defendants who manifest a high degree of concern for others but who, through poor moral luck, are driven to crime as a last resort due to social disadvantage. But it would still leave open the decision of whether to also extend mitigation through other means to defendants who lack sufficient concern for others due to social disadvantage, who would not manifest such avoidance dispositions.

Ultimately, which avenue of reform we should choose may depend on the empirical question of how widespread reluctant purposeful agents are. If it turns out (as I have suggested) that much purposeful wrongdoing is reluctant wrongdoing, so that the exception threatens to swallow the class, it may be simpler to just sort culpable agents by reluctance directly. If, on the other hand, reluctant wrongdoing is a smaller subset, it may be that a mitigating carve-out is more appropriate. This could be done in the structure of a traditional affirmative defense, where the burden of proof lies on the defendant, or, as in the way many jurisdictions treat extreme emotional distress, as a defense where the burden of persuasion lies on the defendant, but where, once met, the burden of proof lies with the state.

There are thus a variety of doctrinal options, each of which provide a different balance for legislatures more concerned with false positives, and false negatives, respectively. What is most important to either kind of reform is that it would provide (in contrast with our current regime) *some* direct avenue for defendants to treat evidence that they pursued numerous alternatives before engaging in wrongdoing as a last resort as grounds for mitigation, even when their purposeful wrongdoing is deliberate and premeditated, and so a poor fit for diminished responsibility excuses.

As a proposal for reform, a system of mens rea that codifies intentional commitments through counterfactual tests has at least three important and related advantages over alternative mens rea reform that would revert to a direct assessment of the defendants' motives or reasons. First, it codifies a more appealing and morally accurate underlying picture of culpability. Second, it identifies a psychological capacity (avoidance commitments) which can be described in normatively neutral terms. Third, it is a psychological capacity which we have every reason to think jurors will be just as accurate in assessing (if not more accurate) as the PKRN states of the current regime.

Crucially, unlike a reform that recasts mens rea regimes in ways that require a general investigation into the presence of malice or motive, the exculpatory mens rea states or excuses proposed here would direct the fact-finder to determine the presence or absence of a psychological state (an intentional commitment) which can be described (and evidenced) in normatively neutral ways. Just as I can evidence a commitment to something utterly normatively mundane (like picking up salmon for dinner at the market) by the fact that I was disposed to keep driving to markets until I found one that sold salmon,[180] a defendant can evidence a commitment to avoiding the taking of a life (or the absence of such a commitment) by whether the defendant was disposed to take alternative means (like trying de-escalation tactics rather than chokeholds).

Unlike other reform efforts described above, this system would not require criminal law to look behind the defendant's more proximate mental states and engage in the normatively messy business of attributing and assessing the defendant's motives. It would allow the legislature to specify which psychological states are culpable or exculpatory (a commitment to avoid wrongdoing) and the jury to determine the factual question of whether that psychological state was present or absent. This could be done by showing that a defendant (like Judy Norman) actually did try numerous alternatives between pursuing criminalized activity as a last resort, or, if possible, by establishing evidence that the defendant was counterfactually disposed to have taken such alternatives had they been available.[181]

More excitingly, not only is this kind of counter-factual analysis a purely factual, rather than normative, inquiry, it is one that is already intimately familiar to the legal system. Juries are already asked to engage in similar counter-factual analysis when asked to determine but-for causation in criminal trials and in many other contexts, such as the 'dangerous proximity test' for attempt liability. Thus, while its application to culpability assessments may be novel, the inquiry itself is one that fact finders are already practiced in engaging in.

Indeed, this kind of dispositional analysis is already a tacit part of the current mens rea legal doctrine. To distinguish intentional action from mere knowing action in terms of tracking dispositions, after all, is to employ just such a counter-factual or dispositional test, asking fact-finders to consider whether the agent would continue to attempt to harm the victim if their current action had failed to produce the resulting harm. What is being proposed here is simply that we replace this particular dispositional analysis with another one which more closely tracks the genuine underlying culpability of the defendants.

This proposed system of sorting mens rea by intentional commitments, evidenced by counter-factual tests, is of course not without difficulties of its own. In particular, just as juries' attributions of intention may be causally influenced by their

---

[180] *See*, SCOTT SHAPIRO, LEGALITY 126-129 (2011).
[181] State v. Norman, 378 S.E.2d 8 (N.C. 1989)

normative frameworks,[182] the way they determine the relevant counterfactuals may be similarly influenced. Indeed, studies into jury attribution of causation suggest as much.[183] The claim here is not that juries would be perfectly competent at applying this test, but rather that the proposed mens rea states are ones that we have every reason to think juries would be just as good, if not better, at applying than the current PKRN states, and easier than asking juries to assess motives.

Articulating at least the broad strokes of a novel mens rea doctrine is comparatively simple because the difficulty lies in the philosophical work of determining what the underlying culpable psychological states really are, so that law-makers can tell the fact-finder to look for those psychological states, rather than ask the fact finder to try to do the philosophical work on their own at trial. The problem is not that we could not craft a legal defense to track these states, but rather that that the current law is crafting excuses under the mistaken assumption that other dispositions (namely, tracking dispositions) are the only ones worth condemning. Once we recognize the existence of avoidance commitments, and that avoidance commitments are just as important to determining culpability, the task of crafting an appropriate set of excuses (or even a new mens rea hierarchy to be applied more generally) is relatively straightforward.

VI. CONCLUSION

In this article I have argued against the traditional view that the criminal law's goal of making criminal liability proportional to the dangerousness or culpability of the defendant supports treating purposeful defendants as more liable than reckless or negligent agents who engage in the same wrongdoing.

On the traditional view, purposeful agents – even those who engage in wrongdoing reluctantly as a means to some further goal – appear more culpable because of certain intentional commitments, constitutive of purposeful wrongdoing, to 'track the harm' to the victim. Philosophical thought experiments such as the trolley problem have purported to show that such tracking commitments exist and are especially dangerous and especially culpable, by considering counterfactual circumstances where the victim might avoid the harm caused by the agent's criminal actions, and showing that the purposeful agent with tracking commitments will continue to engage in wrongful behavior.

This article has shown that if we expand the set of counterfactuals we consider, other overlooked intentional commitments can come into view. In particular, I have argued that a particularly important set of 'avoidance-commitments,' which are present in the case of reluctant purposeful agents but absent in the case of callous agents, speak in favor of *diminished* liability for many purposeful defendants. I have shown how these

---

[182] There is an enormous and growing body of psychological literature on the phenomenon of 'intentionality bias.' For a seminal discussion of how people's intentionality judgments are influenced in particular by the moral valence of the putative intentional action, See especially Joshua Knobe*, Intentional Action and Side Effects in Ordinary Language*, 63 ANALYSIS 190 (2003).

[183] *See, e.g.,* David Lagnado & Tobias Gerstenberg, *Causation in Moral and Legal Reasoning*, *in* OXFORD HANDBOOK OF CAUSAL REASONING 565 (2017).

avoidance commitments provide a promising new class of psychological states for crafting a new mens rea regime which employs counter-factual tests for avoidance commitments to either augment or replace the traditional mens rea classification in terms of the four states of purpose, knowledge, recklessness, and negligence.

In doing so, I hope to have highlighted the importance of how a seemingly neutral mens rea regime might actually function to protect powerful wrongdoers, like police officers who fail to recognize that their victims' lives matter, while further penalizing many of the most marginalized in society – like the survivors of abuse – whose circumstances necessitate the kinds of purposeful choices which others in society enjoy the privilege of never having to face, though we might have been disposed to make the same choices when put in similar circumstances.

The PKRN mens rea regime evidences a commitment on the part of the state toward punishing those who commit purposeful crimes of desperation, while excusing those who commit crimes of convenience, who are unwilling to take easily available options. Failure to be clear-eyed about such commitments creates a further barrier to recognizing the true moral magnitude of failures by police officers to recognize the humanity of those they police.