UNIVERSITY OF CALIFORNIA
SANTA CRUZ

REMORSE AND PUNISHMENT:
APPROACHES, PROBLEMS, AND NEW IDEAS

A dissertation submitted in partial satisfaction
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

In

PHILOSOPHY

by

Travis Raymond

June 2022

The Dissertation of Travis Raymond is approved:

_____
Professor Daniel Guevara

_____
Professor Paul Roth

_____
Professor Nico Orlandi

_____
Peter Biehl
Vice Provost and Dean of Graduate Studies

TABLE OF CONTENTS

REMORSE AND PUNISHMENT:
APPROACHES, PROBLEMS, AND NEW IDEAS

Travis Raymond

Abstract: Prisoners' rights advocates justifiably seek to combat the seemingly ever growing institutions of punishment and increasing incarceration rates in the United States specifically, and the modern state more generally. Any policy with even the modest appearance of potential to reduce punishment is embraced, sometimes uncritically. One such practice is dangling the possibility of less punishment for criminal offenders than they might otherwise deserve if those offenders publicly express remorse for their crimes. I argue that this practice places all criminal offenders under pressure to become remorseful or at the very least express performative remorse. This pressure is enhanced with a threat of violence such that it is theoretically indistinguishable from coercion. Coercion of criminal offenders into adopting a political or moral stance is an assault on their basic human dignity and constitutes a harm that is distinct from punishment itself; one that the state may not justifiably impose. This practice may also have the unintended or unacknowledged effect of increasing the palatability of punishment overall.

From this framework flow contentions about the criminal law more generally, including: the inherently political nature of the criminal law; limitations on its legitimate reach due to that nature; the way the criminal law works to maintain

existing power dynamics and/or mask inequalities; and whether principles of restorative justice might point to a different approach to remorse in the criminal law context.

Finally, I distinguish my approach from a similar call to limit considerations of remorse in the context of punishment. The distinction between sincere remorse and performative remorse has attracted the attention of cognitive scientists who question whether judges and juries are capable of distinguishing the two. Some argue in favor of supplementing or even replacing judges and/or jurors with emerging technology (*e.g.* fMRI examinations) and expert interpretation thereof. I argue these calls are misguided. Concepts like remorse, sincerity, and deception should not be reduced to what is measurable, determinate, and generalizable in the context of the criminal law where the norms are created by communities and are necessarily reinterpreted collectively and subjectively on a case by case basis.

For Hope, Ruby, and Louis
Thank you for your love, support, and patience.

CHAPTER I: INTRODUCTION

> *You will study the wisdom of the past for in a wilderness of conflicting counsels, a trail has there been blazed. You will study the life of mankind, for this is the life you must order, and, to order with wisdom, must know. You will study the precepts of justice, for these are the truths that through you shall come to their hour of triumph. Here is the high emprise, the fine endeavor, the splendid possibility of achievement, to which I summon you and bid you welcome.*
>
> -Justice Cardozo in the 1925 commencement address at Albany Law School

Philosophy and law are natural partners. In the law philosophers find the most systematic recording of humanity's moral dilemmas, uses and abuses of power, and attempts at justice. In other words, philosophers find a treasure trove of real life examples to inspire and ground their theoretical work. In philosophy lawyers and judges find theoretical tools to help make sense of the chaotic mess that is the law. After all, it is the repository of centuries' worth of accidents, popular psychology, fleeting moments of political passion or outrage, good though often misguided intentions, and more. Accordingly, it is fair to question which, if any, of the many byzantine legal practices are justifiable. Philosophy is a most useful tool in that process.

This paper straddles both approaches. First, it is a philosophical interrogation of a specific legal practice: coercing remorse in criminal offenders by permitting adjudicators to impose less punishment when the offender expresses remorse. From that examination, however, flow more general philosophical contentions about the

nature of the criminal law, the proper basis of judgments rendered in criminal cases, and how injustice seeps into the process.

My first job out of law school was as a public defender in Clark County, Nevada covering the Las Vegas Regional Courthouse. Within my first month on the job I was handed a caseload that included more than one hundred clients charged with various misdemeanors and felonies. Following convictions or plea deals I was ethically bound to inform my clients that if they apologized and expressed remorse when they were allowed to speak on their own behalf at their sentencing hearing the judge was permitted to impose less punishment than the judge otherwise thought they deserved. At first, I found nothing remarkable about this practice. But over time, as nearly every client of mine took advantage of this opportunity, even in cases where I was unsure they had any reason to feel sorry (*e.g.* minor drug possession), I started to suspect something more sinister was going on. My pursuit of academic philosophy generally, and certainly this paper specifically, is an attempt to scratch this particular itch. In Chapter II I use philosophy to explain what I find so troubling about the current role remorse plays in the criminal justice system: the state is using violence to coerce offenders into sincerely altering their political/moral beliefs in infringement of a basic right to autonomy or, at the very least, forcing offenders into performative expressions of political/moral beliefs in line with the dominant paradigm.

Chapter III explores the importance of political beliefs and attitudes in the context of the criminal law by exploring what it might mean to say that all crime is political, all prisoners are political prisoners, and that there is a right to immorality (at

least in thought). Criminal law is incongruent with morality and is influenced to a significant extent by existing power dynamics within society. Commentators who focus only on the moral dimensions of a criminal act or practice are missing the important ways this political dimension should shape the way we view both individual obligations and rights, as well as constraints on the state. Developments in ideal vs. non-ideal theory suggest that many ordinary crimes ought to be viewed as a sort of political resistance in societies where the existing power dynamics subject portions of the population to various forms of injustice. Moreover, acknowledgment of the role of interpretation in legal practice may require expanding this view even beyond societies where there is significant injustice. The upshot is that while a given state may be justified in punishing certain actions, it is not justified in policing the morality of thought by imposing more violence on offenders who refuse to express remorse.

If I am right and the imposition of less punishment when criminal offenders express remorse does not serve the criminal law's proper aim of justice, then it is fair to wonder why the practice has such intuitive appeal. What function is it serving for society and the various actors in the legal process who participate in it? In Chapter IV, I explore these topics via a philosophical genealogy in the spirit of Rousseau, Nietzsche, and Foucault. I trace the role of remorse from its most natural context in private relationships, through its warped transformation into a political tool by the Catholic Church during the Inquisition, before arriving in our current criminal codes and courtrooms without the explicit dogma of oppression but no less useful to the

preservation of existing power structures and dynamics. It may be tempting to view extensions of mercy due to expressions of remorse as an expression of the virtues of forgiveness or magnanimity, but I argue that this may be a mask for darker retributive attitudes and emotions. I try to articulate what a more just version of remorse in the public sphere might look like via trends in restorative justice.

In Chapter V, I distinguish my approach from a similar call to limit considerations of remorse in the context of punishment. The distinction between sincere remorse and performative remorse is attracting attention from a wide variety of academic commentators, including those typically outside of philosophy and the law. Cognitive scientists are questioning whether judges are capable of distinguishing sincere expressions of remorse from opportunistic faking. With varying levels of urgency they argue in favor of supplementing or even replacing judges and/or jurors with emerging technology (*e.g.* fMRI examinations) and expert interpretation thereof. I examine the approach of these scientists as well as some philosophical theories in support of their approach. I argue they are misguided. These commentators reduce concepts like remorse, sincerity, and deception to what is measurable, determinate, and generalized. I argue that this approach is at odds with the criminal law where the norms are created by communities and are necessarily reinterpreted collectively and subjectively on a case by case basis.

Finally, in Chapter VI I conclude the paper with a short dialogue between a philosopher, scientist, and lawyer that anticipates criticism and acknowledges tension in the paper and my own thinking.

CHAPTER II: CRIMINAL LAW AND COERCED REMORSE

> *Not merely to extract your confession nor to punish you. Shall I tell you why we have brought you here? To cure you! To make you sane!…The Party is not interested in the overt act: the thought is all we care about."*
>
> -George Orwell 1984

Almost every day a familiar scene plays out thousands of times in the United States: an individual convicted of a criminal offense stands before the court awaiting the determination of the extent and scope of their punishment. The offender is given an opportunity to speak before the sentence is rendered and, more often than not, they begin to apologize. They apologize to the victim (if there is, in fact, a person they have directly harmed), to the community, to the arresting officers and to the judge for having to perform their duty. If they are well-counseled they appeal to the personal interests/preferences (religion, commitment to family, etc.) of the decision-maker. They speak of remorse and repentance; how they wish they could take back their crime, how they wish they could make up for it in some way. Then, finally, they plead for the state's mercy. Conspicuously absent from many of these proceedings is a discussion of the crime itself.

The cynical observer will believe the apologies are insincere, the proclamations of remorse and repentance exaggerated or outright lies. But with the full power of the state conspicuously arrayed against the individual, who is often already in chains and monitored closely by armed guards, the question becomes not whether the offender is proffering sincere remorse and repentance, but how could

they possibly refuse?[1] This paper argues that the state unjustly coerces offenders into remorse via the possibility of mercy from state punishment. In particular, it is an illegitimate use of the state's power to coerce individuals into specific moral and political beliefs or attitudes.

A. Distinguishing mercy from considerations of blameworthiness.

The aim of this paper is to question the moral worth of the state's imposition of mercy in criminal cases because a criminal offender is remorseful. In order to proceed we must have a clear understanding of what mercy as state action consists of. Jeffrie G. Murphy has written extensively on the topic and his definition is a good starting point: mercy is the imposition of a less harsh consequence on a wrongdoer than is permitted by institutional or legal rules.[2] A paradigmatic case of mercy would be for a sentencing judge[3] to sentence a convicted, cold-blooded murderer to life in prison in a state that permits the institution of the death penalty.[4]

Murphy's definition distinguishes mercy from considerations of

[1] Later in this chapter I acknowledge that coercing actual remorse and insincere expressions of remorse are distinct, but I maintain that the offender has suffered a harm in either case that it is impermissible for the state to inflict.

[2] Murphy 2003: 13-14

[3] Or a jury, as often occurs in capital cases.

[4] Besides sentencing a criminal to lesser punishment than permitted following a trial or guilty plea, there are many other state actions that may constitute mercy under the proposed definition including, but not limited to: the commutation of a death sentence to life in prison by a governor; not imposing the full prison term an offender is sentenced to by granting parole through a parole board; pardoning an offender; etc. For clarity and continuity's sake this paper will primarily reference mercy at the sentencing stage of the criminal process, but the criticism of mercy as state action applies to every circumstance fitting the definition.

blameworthiness, which conforms to the way we generally speak of mercy. For instance, a state that refrains from punishing individuals who are innocent even though the state is capable of obtaining lawful convictions is not said to be merciful. Rather, a state that does not punish the innocent though it is capable of doing so in a lawful manner is called "just" and not "merciful." A state that contravened the maxim would be called "unjust." Thus, the term mercy is properly applied only in those situations where a state may justly punish an offender to some extent, but imposes less punishment than it is entitled to. This distinction is important because there are two situations that commonly arise in criminal matters and are often confused with mercy: excuse and justification.

An excuse occurs when an individual engages in conduct that is wrong but the individual cannot be blamed because the individual was not a responsible agent.[5] The paradigmatic excuse is insanity. The paradigmatic case is something like: Jones kills Smith by shooting him with a gun but Jones is operating under the genuine delusion that she is healing him with a magic wand. To argue that Jones should receive a lesser punishment than a cold-blooded murderer (or no punishment at all) is not to argue that Jones should be granted mercy. Rather, the argument is that Jones is less (or not at all) blameworthy. In this hypothetical, Jones' lacked the necessary *mens rea* to be blameworthy.

Justification is theoretically distinct from excuse, but operates in a similar

[5] Kadish & Schulhofer 2001: 842-950

fashion. A justification occurs when an individual engages in conduct that is considered wrongful in most instances, but is considered just when all of the circumstances of the particular instance are weighed.[6] The paradigmatic justification is self-defense. The paradigmatic case is: Jones kills Smith by shooting him with a gun after Smith has cornered her and approached with a knife saying, "I'm going to kill you." As with an excuse, Smith does not qualify for mercy because Smith is not blameworthy. Returning to our scene in the sentencing courtroom, a judge who imposes a less harsh sentence than permitted on a criminal offender because the judge believes there to be credible evidence of an excuse or justification is not committing an act of mercy, but trying to do justice under the maxim that only the blameworthy should be punished.[7]

The foregoing analysis of excuse and justification can be applied to any considerations of an offender's crime itself or to characteristics of the specific offender in existence at the time of the crime. Consider two thieves: one who is rich

[6] Kadish & Schulhofer 2001: 750-842

[7] Carla Ann Hage Johnson (1991) has argued that by excluding from the definition of mercy a judge's or pardoner's attempts to do justice by weighing individual cases Murphy has "ignore[d] the message of history" under which such discretion evolved. The argument deserves more consideration than a footnote permits, but as the primary purpose of this paper is not to eliminate "mercy" from the legal lexicon but to argue for the elimination of one specific form of mercy (i.e. granted on the basis of remorse or repentance), a footnote is all the space that can be spared. I think that there is evidence elsewhere in the history of criminal law that proves that such discretion relates to blameworthiness and not mercy. For instance, murder is separated into degrees on the basis of the intent of the killer at the time of the killing and a judge's discretion in punishment for murder is determined by the degree an offender is convicted of. And the differing degrees of murder correspond roughly to the degree of blameworthiness for the killing.

and well-fed and the other poor and starving. They are both convicted of stealing identical loaves of bread and appear before a sentencing judge. If the sentencing judge imposes a less harsh sentence on the starving thief it may be tempting to describe the act as merciful. However, for most people, it is their intuition that the starving thief is entitled to less punishment. There are many possible reasons for this intuition. We may conceive of starvation to be a sort of excuse (i.e. because they were starving the thief could not resist taking bread and therefore is less responsible for the crime than someone not under the duress of starvation). Starvation could also be thought of as a justification (i.e. that we believe it is wrong to steal food, but considering that the thief was starving and it was necessary to eat, it was not wrong in this circumstance). Or we may simply hold that any law (or legal system) mandating starvation in any case is *prima facie* unjust. On any of these bases it is evident that we believe the starving thief is entitled to less punishment because we believe that they are less blameworthy.

To summarize the argument thus far: it is permissible for the state to consider the circumstances of the crime and the particular characteristics of the offender in existence at the time of the crime in order to determine blameworthiness. Indeed these considerations are necessary to achieve justice, as only the blameworthy should be punished. But these lesser punishments are not properly understood as mercy because mercy is the imposition of a lesser punishment than the offender deserves due to their blameworthiness, and such considerations establish that the offender was somehow less blameworthy and therefore entitled to less punishment as a matter of

justice.

Now let us consider a third thief: one that was not hungry at the time she stole the loaf of bread, but she is starving at the time of her sentencing hearing (through no fault of the state). This hunger is a characteristic of the thief that did not exist at the time of the crime and has no bearing on her blameworthiness for that crime. Yet, modern criminal systems generally permit the state to consider such characteristics and impose less punishment.[8] Thus, it seems, Murphy's definition requires augmentation. A complete definition of mercy as state action is the imposition of less punishment on a criminal offender than is permitted by law and justice due to characteristics of the offender that do not relate to the offender's blameworthiness for the crime committed.[9] Chief among such characteristics that states consider in granting mercy is whether the criminal offender expresses remorse.

B. Remorse is distinct from typical considerations of blameworthiness in the context of the criminal law.

John Deigh contends that remorse is dependent on a conception of value. Remorse is experienced when one has damaged or destroyed an object of value and is

[8] Some other examples here include whether the offender is aged and/or infirm, whether the punishment will pose an undue burden on the offender's family, or the victim, or society, etc. I do not claim that there is anything problematic with mercy in these cases. It is solely cases where the possibility of mercy infringes on the autonomy of individuals that I am concerned with in this paper.

[9] To be fair, Murphy may have anticipated such an addition to his definition of mercy as he states in his earlier writings "It strikes me as analytic that mercy is based on a compassionate concern for the *defendant's* plight." Though in this context Murphy is specifically arguing that the imposition of less punishment due to concerns of utility is not actual mercy. Murphy & Hampton 1988: 173.

incapable of providing a remedy.[10] To define repentance, Murphy builds from Deigh's definition of remorse. Repentance is similar to remorse, but it includes a social dimension. For Murphy, repentance is not just an interior mental act of remorse, but a desire or commitment to make amends, possibly to the community as a whole.[11] Additionally, remorse and repentance carry a moral judgment that one's actions and "the aspects of one's character that caused the actions" are wrong, perhaps even evil, and must be repudiated.

It may be tempting to argue that remorseful and/or repentant criminal offenders are less blameworthy for their crimes than the remorseless and unrepentant, but there are good reasons to refrain from doing so. There is a clear temporal element to both remorse and repentance. True remorse and repentance occur only after an act of wrongdoing as they are mental responses to the harm one has caused. In contrast, we generally tend to feel resentment toward people who cause harm due to their intent at the time of the action. As Peter Strawson put it, "If someone treads on my hand accidentally, while trying to help me, the pain may be no less acute than if he treads on it in contemptuous disregard of my existence or with a malevolent wish to injure me. But I shall generally feel in the second case a kind and degree of

[10] Deigh 1996.

[11] Murphy 2003: 41. I return to this distinction between remorse and repentance in Chapter IV where I think it is useful as part of tracing how remorse went from being a private matter to something that the state had an interest in during the Spanish Inquisition. In this chapter, I use remorse primarily because the distinction is unimportant to the argument I make. Coercing remorse or repentance is unjust.

resentment that I shall not feel in the first."[12] While we may choose to forgive an individual for causing harm based on their subsequent mental state, whether we deem them worthy of blame in the first place is based entirely on their mental state at the time of the action that caused the harm.

Additional evidence that remorse and repentance are not indicators of blameworthiness is that they may justifiably be experienced by one who has done nothing wrong but has caused harm nonetheless. The driver of an automobile who suffers a heart attack that they had no reason to suspect was imminent, loses control of their vehicle, and injures or kills another person may still justifiably feel remorse. They may even repudiate any personal habits that increased their likelihood for heart attack. As Deigh suggests, remorse is evidence of a general ethical orientation of care for that which is valuable.[13] In a sense, remorse may be evidence of a good character, but not an indication of blameworthiness for any individual act.

Given that remorse and repentance are not indicators of blameworthiness for a specific crime it is puzzling why the state considers them at all in determining a just punishment for that crime.[14] However, modern states are doing more than just

---

[12] Strawson 1982: 65.

[13] Deigh 1996: 52.

[14] This paper approaches the problem of mercy and coerced remorse through a retributivist framework. Of course, a consequentialist might think there are still good reasons for imposing mercy on the basis of remorse or repentance even though they are not related to the blameworthiness of an offender. For instance, if expressions of remorse were correlated with lower rates of recidivism. A longer response may be warranted, but for now I would just respond that if the rest of the paper is convincing then at the least the consequentialist in their calculations/weighing of the consequences will have to account for the significant harm suffered by offenders

considering remorse as they impose punishments, they are coercing criminal offenders into these mental states through the extension of the possibility of mercy.

C. Mercy by the state based on remorse amounts to coercion of criminal offenders to be remorseful.

Coercion is generally defined as the use of force or threats of force to pressure another into taking (or refraining from) some action. There are two obstacles to establishing that modern states are coercing criminal offenders into remorse and repentance: 1) whether modern punishment is force sufficient to coerce; and 2) whether permitting less punishment for the remorseful and repentant constitutes a threat to the remorseless and unrepentant.[15]

The first hurdle is the easiest to overcome. While there may be certain segments of the population that believe modern punishment is not harsh enough on criminal offenders, as a theoretical and practical matter, modern punishment is almost certainly severe enough that rational actors should wish to avoid it. From a

---

coerced in the way I'm suggesting. Moreover, if the right to be free of coerced remorse is as strong as the right of, say, the innocent not to be punished, then consequences like deterrence and a reduction of recidivism may not matter.

[15] One might suggest a third obstacle: whether the state has the intent to coerce or whether any resulting coercion is the byproduct of some other purpose. In a sense it is impossible to declare the intent of the modern, often democratic, state, where numerous different factions may support the same policy for wholly distinct reasons. The best that can be said is that at least some factions view the imposition of mercy as a means to foster remorse and repentance. Murphy (2003: 112-113) cites the *Evangelium Vitae* of the Catholic Church in support of such a view. Additionally, I think that it is inherent in the modern movement for "rehabilitation" that criminal offenders are intended to undergo some psychological change. Finally, I posit some structural causes of the state's coercion of repentance and remorse in Section E of this chapter that suggest remorse and repentance serve to bolster the state.

theoretical perspective, H.L.A. Hart and others have defined punishment such that "It must involve pain or other consequences normally considered unpleasant."[16] Thus, anything less cannot legitimately be considered punishment and the withholding of such treatment cannot legitimately be considered mercy. From the practical perspective, it can be plausibly argued that modern prison sentences and the practice of probation as a form of punishment are less painful than the more classical corporal punishments and, therefore, less able to coerce. However, the vast majority of the evidence from modern criminology and popular culture[17] have proven significant psychological and physical harm resulting from modern prison and probationary systems, such that any rational actor would obviously try to avoid them. It would be a rare exception for a rational criminal offender to not prefer a 5 year prison sentence to a 20 year sentence, probation to prison, and no sentence at all to probation. There can be little doubt that modern forms of punishment constitute sufficient force to coerce.

Establishing that mercy for the remorseful or repentant constitutes a threat to the remorseless or unrepentant is slightly more difficult, but only because the threat is often implicit. It is true that one will rarely hear the numerous state representatives involved in the imposition of punishment directly threaten a remorseless or

[16] Hart 1968 and citations therein.

[17] I am referencing here everything from Foucault (1995) to modern scientific studies of the psychological impact of punishment (Haney 2008; Haney & Zimbardo 1998), and even the widespread fear of punishment encouraged by pop culture works like MSNBC's television program "Lockup."

unrepentant offender with harsher punishment unless they show remorse. However, it does happen on occasion. In 2020, during the height of pandemic stay-at-home orders a salon owner in Dallas, Texas refused to shut down her business and participated prominently in protests against the orders. Shelley Luther was found in violation of state and local health orders. During a hearing to determine the consequences of her violation the trial judge told Luther that she had the "keys" to the county jailhouse. "You may utilize them now if you would like to take this opportunity to acknowledge that your actions were selfish, putting your interests ahead of those of the community in which you live." The court informed Luther that if she would publicly express contrition then, "This court will consider the payment of a fine in lieu of the incarceration, which you have demonstrated you have so clearly earned." Luther refused to apologize or show remorse stating: "I have to disagree with you, sir, when you say I'm selfish, because feeding my kids is not selfish. I have hairstylists that are going hungry because they'd rather feed their kids. So, sir, if you think the law is more important than kids being fed, then go ahead with your decision, because I'm not going to shut the salon." The judge immediately remanded Luther into custody and sentenced her to seven days in jail.[18]

While the express threats detailed in that case are unusual, the threat itself hangs over the head of every criminal offender. It is codified in most modern criminal statutory systems and routinely put into practice. The codification takes the form of statutes defining remorse as a mitigating factor and remorselessness as an

[18] In re Shelley Luther, No. 20-0363 (Tex. 2021).

aggravating factor that sentencing judges and parole advisory boards may consider in their discretion.[19] Additionally, the state's implicit threat often becomes explicit in practice, through the offender's attorney.[20] Attorneys are ethically bound to inform the criminal offender that his/her remorse and repentance may result in less punishment, thereby directly communicating the state's threat that if they do not become remorseful and repentant they are more likely to receive greater punishment. By granting remorseful and repentant offenders less punishment (or at least the possibility of less punishment) the state places pressure through the threat of force on remorseless offenders to, in some sense, change their mind.

The contention may arise that the state is only pressuring offenders to insincerely appear remorseful or repentant ("performative remorse"), and not to adopt a genuine state of mind. However, this contention is most meaningful if one assumes that there is no way to reliably detect sincere remorse and repentance. If that is truly the case, then the state's practice of granting mercy based on remorse and repentance is of questionable moral worth indeed.[21] Moreover, if the state is infringing on

[19] In some statutes remorse is mentioned by name (*California Penal Code sec. 502*), but even where there is not a clear statutory basis for consideration of remorse courts have created a basis out of the common law (e.g. *Echavarria v. State*, 839 P.2d 589 (Nev. 1992)(per curiam)).

[20] It is arguable that at least some, if not all, criminal defense attorneys operate in the capacity of state actors. Public defenders, for example, are often direct employees of the state. And all attorneys are supervised by state administrative offices, such as bar associations.

[21] The distinction between sincere remorse vs. performative remorse and the proper criteria for judging so are the subject of Chapter V.

human dignity by coercing sincere remorse, as I argue in the next section, then injustice may occur even where the state only coerces performative remorse.

D. Is it just to coerce offenders to be remorseful?

It is my contention that basic respect for human dignity precludes coercing individuals to adopt beliefs, especially moral and political judgments of the sort that are inherent in remorse. In arguing for this position I will borrow from work by Jon Stuart Mill, Gerald Dworkin, and Herbert Morris. That human beings have a right to be free from what I will call mental coercion is founded on respect for the autonomy and rationality of human beings. It is also my contention that such a right is not negated by an act of criminal wrongdoing for which an individual may be justifiably punished.

1. Political Prisoners

I will begin where I believe common intuition is most strongly in support of my position: political prisoners. Murphy, who ultimately argues that genuine repentance may be a justified ground for mercy as state action, concedes that we would not want to see repentance from persons who commit minor crimes from a motive of nonviolent civil disobedience. Murphy asks, "Do we really want to seek repentance from the Martin Luther King Jr.s and Gandhis of the world?"[22] However, Murphy's analysis of why we would not wish to encourage remorse is superficial. Murphy states "contemporary criminal law (at least in America) tends toward radical

[22] Murphy 2003: 46.

over criminalization--punishing many offenses with absurd excess and regarding some actions as crimes that, since their moral wrongness is doubtful, are also doubtful objects of repentance."[23]

I believe Murphy's analysis is missing something crucial about our intuition in these cases. We do not wish to see political prisoners coerced by the prospect of punishment into repentance not just because we, sometimes, believe their actions were justified and not crimes at all. We are intuitively opposed to mental coercion on a deeper level because we believe that the use of force to coerce individuals to renounce their personal beliefs and to denounce themselves is inhumane. To illustrate the point, we must consider another type of political prisoner; one whose crimes are not minor and whose political ends we often cannot agree with: political terrorists.

In 1995,Timothy McVeigh detonated a truck bomb in front of the Alfred P. Murrah Federal Building killing 168 people, as revenge for "what the U.S. government did at Waco and Ruby Ridge." Prior to his execution in 2001, McVeigh never expressed remorse nor repentance, and even went so far as to claim that he would have given serious consideration to using sniper attacks in a war of attrition against the government if exposed to the idea prior to the Oklahoma City bombing.[24] McVeigh was executed by the state, the harshest punishment U.S. federal law permits, a punishment that many find justified. Very few would claim that

[23] Murphy 2003: 45-46.

[24] Michel & Herbeck 2001: 304.

McVeigh's actions and the intent with which he took them were not immoral, blameworthy, and justification for punishment. There are also few that agree with McVeigh's political position that the government has no right to interfere with a heavily armed militia that may be abusing children. And yet my sense is that most people would not agree that threatening McVeigh with more painful forms of punishment would be justified to coerce McVeigh into experiencing remorse or repentance; coerced into a mental state where he believed his actions morally wrong and his character flawed or even evil.

If I am correct and others share my intuition, our common intuition is likely not based on any belief that McVeigh's crime was justified or that his political motivations had any merit. Nor is it based on a belief in American "over criminalization." The source must not be something specific to McVeigh or his circumstances, but more universal.

2. What is offensive in mental coercion is infringing on the autonomy and rationality of criminal offenders.

In his famous paper "Paternalism," Gerald Dworkin parses a passage from John Stuart Mill's *On Liberty* wherein Mill argues that an individual's own good is never a sufficient warrant for interfering with the liberty of action of that individual.[25] Dworkin argues that Mill is not relying solely on a traditional utilitarian basis for this contention, because Mill is uncharacteristically absolute in this proscription against paternalistic infringements on liberty. If Mill's argument is only that individuals

[25] Dworkin 1971.

know their own interest better than others or society do, then it is at best a contingent argument where the context and circumstances of the individuals will have to be considered, but might be outweighed by other considerations.

Rather, Dworkin claims that the absolute nature of the prohibition on paternalism is tied to a non-contingent argument about human autonomy and rationality:

> When Mill states that "there is a part of the life of every person who has come to years of discretion, within which the individuality of that person ought to reign uncontrolled either by any other person or by the public collectively," he is saying something about what it means to be a person, an autonomous agent. It is because coercing a person for his own good denies this status as an independent entity that Mill objects to it so strongly and in such absolute terms. To be able to choose is a good that is independent of the wisdom of what is chosen. A man's "mode of laying out his existence is the best, not because it is the best in itself, but because it is his own mode." It is the privilege and proper condition of a human being, arrived at the maturity of his faculties, to use and interpret experience in his own way."[26]

I contend that respect for human dignity, i.e. autonomy and rationality, prohibits the mental coercion of criminal defendants. Even when we disagree with the wisdom of a defendant's judgments and/or beliefs, they are entitled to respect. While all coercion is an affront to human dignity, there are two strong reasons for

[26] I acknowledge that there are other, strictly consequentialist, interpretations of Mill. I cite Dworkin's here primarily because it forcefully and succinctly captures what I contend is wrong with the state coercing certain beliefs. I also happen to think Dworkin's interpretation is an insightful/appropriate interpretation of Mill, though I do not defend that claim here; nor is it essential to the claims of the paper.

suspecting that interfering with autonomy and rationality by coercing remorse is especially troubling.

First, remorse is inextricably bound up with moral and (in the context of the criminal law) political judgments. To express remorse in the context of the criminal law is to affirm the justice/morality of the system, its institutions, and laws. Perhaps the reason why our intuition is so strong in the context of political prisoners, if Murphy and I are right about the strength of intuitions there, is that in these cases it is obvious that the state is using threat of force to compel individuals to adopt certain moral and political beliefs or attitudes. In broad terms, ever since Hobbes the dominant position in political theory is that it is the right of individuals to develop their own conception of the good life. And the state's power is constrained such that it may legitimately exercise power to maintain the conditions which facilitate this pursuit by individuals but not dictate the direction or conclusions of that pursuit.

With this background in mind, it can be said that the harm occurs whether the coerced expression of remorse is genuinely felt by the offender or is purely performative. For surely the right to freely and rationally form our own opinions on the very most important questions (justice, morality, etc.) surely must also entail the right to freely and peacefully express those opinions without threat of violence. Otherwise, it is a very hollow right. One that would force citizens into wearing a hideous public mask that supports the dominant paradigm in all but their most private/intimate moments. For this reason, I do not find a meaningful distinction in the harm suffered by criminal offenders whether they are coerced into sincere

remorse or merely performative remorse.

Additionally, as I introduce in Section E below and articulate in detail in Chapter IV, public expressions by criminal offenders in support of the morality/justness of the law may have the concerning practical consequences stifling dissent, emboldening individuals to punish offenders, and increasing the amount of punishment in society overall. Thereby, the dominant moral and/or political paradigm is entrenched even further.

Second, in contrast to coercing someone into some physical act (robbing a bank against their will or signing a contract under duress) coercing a person into a particular mental state can interfere with features at the core of their individuality. The epigraph to this paper is intended to call to mind Orwell's fictitious state and its ability to make people stop loving each other. In the novel, the state is able to coerce a citizen, Winston, to not only stop an illicit affair, but to not even love or desire his partner, Julia, any more. Winston is so sympathetic not merely because the reader believes the state has deprived him of the future enjoyment of a life with Julia, but because the state has fundamentally altered who he is against his will. It has manipulated his explicit thoughts, implicit associations, emotions, and values. In essence, it has altered his very soul. Moral and political beliefs are almost certainly fundamental to our identity in a similar way to our closest relationships. For instance, remorse and repentance have value in some religions, and to subvert the will of criminal offenders in these situations may be to infringe on an offender's right to salvation through becoming repentant via an independent act of will. On this

understanding, the coercion of criminal offenders to become remorseful and repentant is a particularly egregious assault on their autonomy.

3. Extrapolating beyond political prisoners.

There are three reasons I believe the foregoing analysis is not limited to crimes with a clear political motive nor to criminal offenders that leave behind notes detailing their political motivations, as McVeigh did. First, there is a sense in which all crime is political; that all criminal statutes are the expression of political power; and that all criminals are in rebellion against the sovereign. This is not to say that there is not a distinct moral dimension to the criminal law as well. But it is to say that the political dimension is always present; that we cannot properly conceptualize what it is just or legitimate to do in the context of the criminal law by attending solely to morality.[27]

Here are some examples. As implied *supra*, criminal laws that protect property rights to the extent that a person is prohibited from stealing even upon starvation express a political position; namely a very strict and inhumane conception of individualism. Even crimes that appear at first glance to be the most personal and least political, such as rape and domestic violence, are expressions of political power. Recent trends toward stronger penalties and the revision of available defenses in relation to these crimes have likely been caused by the increase in political power of women, the most common victims of these crimes. And very important work has

[27] This claim is explained and argued for in Chapter III.

been done to make salient just how much of the criminal law and its enforcement practices are efforts to preserve the inferior status of African Americans following the end of slavery and the repeal of formal segregation laws.[28] Individuals that commit a crime can be understood to have rejected the political position embodied in those statutes, with varying levels of intention. Thus, in a very real sense there are only political crimes and political prisoners.

Second, it is not merely the coerced renouncement of one's political beliefs and realignment with the dominant political ideology that we might be intuitively opposed to. It is the coerced renouncement of one's own character and choices freely made based on that character. It is being coerced into a more general moral position than most political stances take the form of.

Finally, even if one were to believe that a particular criminal statute perfectly captured objective moral norms, sometimes referred to as "natural duties," one might still believe that the state is not justified in using force to compel agreement with these norms. Respect for autonomy on the grounds I've argued for above might entail something like a political right to hold immoral beliefs, as I argue in the next chapter.

4. The fact the state is entitled to punish does not render the mental coercion permissible.

William Edmundson has argued that if the state is justified in punishing

[28] Alexander 2012.

wrongdoers, then threatening to do so does not make criminal offenders any worse off than they ought to be.[29] The argument works only so long as the act of threatening does not create harm beyond the scope of the punishment contained in the threat. Suppose there is a crime punishable by up to ten years in prison and that such punishment is just; and that the sentencing judge has the discretion to impose that full just punishment or to sentence the prisoner to as little as one year in prison. A remorseless and unrepentant offender sentenced to the full ten years for such a crime will almost certainly experience corollaries to the punishment that may increase its harm, such as concern that the offender will miss key moments in her loved ones' lives, though these are largely beyond the state's control. In contrast, when the state threatens the remorseless and unrepentant offender with the same ten year sentence or the possibility of some lesser punishment if the offender will repent, the state is responsible for both supplementing the harm the offender will suffer and changing its nature.

Instead of merely facing the decision to commit the crime and suffer ten years imprisonment and all the negative corollaries, there are now three possible scenarios. In the first, the offender acquiesces but receives the same ten year prison sentence. Clearly this offender has suffered additional harm. The offender is receiving the maximum just sentence and has had her dignity violated by being coerced into renouncing herself. In the second scenario, the offender acquiesces and receives a

[29] Edmundson 1995.

sentence of less than ten years. In this scenario the state has changed the nature of the harm that the offender will endure. While ten years in prison is just punishment, it is unclear that having one's dignity assaulted is comparable to a prison term. In the third scenario, the offender refuses to repent and receives the full ten year sentence. This offender is serving a just sentence *per se*, but the state is now responsible for additional circumstances that cause additional harm, such as the experience of regret for not having acceded to the coercion. The very act of being threatened may carry additional harm that makes it unjust, even if the content of that threat is no more than what the state may justifiably impose on criminal offenders.

5. The fact that offenders themselves likely would prefer to have the possibility of mercy based on remorse does not render the practice permissible.

One might object to my argument by claiming that if they were accused of a crime, they would want the option of expressing remorse for the possibility of less punishment. They might argue that they personally would prefer the harm of infringement of their autonomy over the harsh treatment of punishment. Or even that such a preference is evidence that no coercion is occurring at all.

Considering the criminal prohibition on blackmail may be useful in responding to this imagined interlocutor. At first glance, blackmail is puzzling. "[B]lackmail appears to consist of the very things that render commercial exchange *non*criminal: blackmail is the act of obtaining or attempting to obtain something of value from persons by offering in return to do something that is noncriminal (i.e., withhold incriminating or embarrassing information) under the threat of otherwise

doing something that is also noncriminal (i.e., disclose information that can be lawfully commodified and sold for the actor's personal gain)."[30] But I think that my argument above illustrates why we might think that blackmail is wrong. The blackmailer is entitled to harm the person in one particular way, but not another. The blackmailer is entitled to publish the information and cause the person embarrassment. And yet this harm does not infringe on the victim's autonomy. However, by presenting the victim with the threat of blackmail, the blackmailer is causing an altogether different kind of harm. The victim of blackmail is wrongfully placed into a position where they must weigh their values against each other. They are ostensibly "choosing," but they lose out no matter which option they take because they will be confronted with an involuntary reordering of their values. They will have been forced into the knowledge that they value money more than their own pride or vice versa. The harm is there no matter which choice they make.

Returning to the objection at hand, just as it does not matter to our calculations of whether the blackmailer has done something wrong that the blackmail victim preferred loss of money to embarrassment, it does not matter in our assessment of whether the state is acting wrongfully that offenders might prefer infringements of their autonomy to incarceration.

My argument has not been that harm of coerced remorse is somehow worse than the harm inherent in punishment, it's that it is impermissible for the state to inflict one type of harm but not another. I will develop that argument further in the

[30] Westen 2012: 595.

next section.

6. How coercing remorse and repentance differs from punishment.

One might contend that punishment itself is an infringement on autonomy and ask why my arguments are targeted only at one specific practice rather than the entire institution. In this section I will contend that the imposition of punishment respects autonomy in a way coerced remorse and repentance cannot.

One way to distinguish punishment from coerced remorse is to notice that the aim of punishment is much more reserved. The goal of punishment might be understood to be ensuring minimum requirements of behavior are met. In contrast, coerced remorse cannot be understood in this way. The target of a policy or practice of coercing remorse must be something like the offender's character. In this way, these policies correspond to what Herbert Morris calls guilt based morality and shame based morality (punishment to guilt and remorse to shame).[31] Guilt based morality is primarily concerned with ensuring that people meet minimum standards, whereas a shame based morality blames people for failing to meet an ideal standard. Morris offers a number of reasons why guilt based morality is an appropriate model for the criminal law and a shame based model is not, but there are two that are particularly relevant to autonomy. First, shame based morality is not essentially tied to fault. As discussed above, we might feel remorseful or repentant for things that we did not choose. Whereas it seems clear that the state may only legitimately punish those who are at fault for their actions. Second, one cannot divest oneself from shame unless

[31] Morris 1971.

one rises to meet the ideal, whereas punishment divests oneself of guilt ("the offender has discharged their debt to society"). The point here is that punishment is an intrusion on autonomy, but the degree is much less than in a system that includes coerced remorse.

Drawing from Morris' work, there is another way to draw a contrast between punishment and coerced remorse. Morris contends that punishment is consistent with respect for autonomy.[32] Morris sketches an idealized account of the criminal law in defense of his position. Autonomous agents can agree to bear burdens and benefits equitably. They can agree that a state enforce this distribution. Accordingly, punishment can be understood as the state imposing a burden on one who has unjustly or impermissibly taken a benefit they weren't entitled to in order to redistribute benefits and burdens in accordance with the agent's initial agreement. In a way, we can say this agent has literally chosen to be punished by assenting to the system and by taking some action to obtain a benefit (or shirk a burden) impermissibly. In this sense, punishment is an affirmation of this agent's autonomy. It respects the agent's choice in how to best arrange society and their choice about how to act within that society.

It is difficult to insert coerced remorse or repentance into this model. There is no problem in imagining that agents might choose a system that coerces remorse (to a degree, we have), and that enforcing such a policy respects that choice. But coercing

[32] Morris 1976.

remorse does not seem to honor the choices of agents' about how to act within that society. In a way, coercing remorse is a formal method of preventing one from ever changing their mind.

E. Structural explanations for coerced remorse.

This chapter has focused on the narrow question of whether the possibility of mercy on the basis of remorse or repentance places unjust coercive pressure on criminal offenders. Aside from arguing that mercy is not tied directly to blameworthiness, this paper has largely remained silent on the question of whether remorse and/or repentance generally have moral worth outside of the context of state action. Chapter IV addresses this topic.

However, here I would like to introduce a structural explanation for the prevalence of the practice. In short, that explanation is that the practice empowers those very same institutions.

It is against the majority of people's inclination to cause another person to suffer, and suffering is inherent in punishment. This inclination may be overcome when we experience the retributive emotions, such as resentment and contempt. However, we are only likely to experience the retributive emotions when we ourselves are the direct victim of an offender or when that offender engages in sufficiently heinous conduct that our retributive emotions are vicariously stoked. The vast majority of criminal offenses that the modern state punishes do not meet this criteria. Picture drug offenses and petty theft. In most cases, there will be very little retributive will toward the criminal offender.

There is another way to overcome our reluctance to punish. We may be more willing to inflict the harm inherent in punishment on offenders if the offenders themselves request that we do so. It is difficult to sentence an individual guilty of mere drug possession to incarceration, but it gets easier if we first listen to the offender impugn his own character, state that he is blameworthy and wants to atone for his wrongful action. Discussions of the specific characteristics of the offender's actual crime, blameworthiness, and suitability for punishment are unnecessary when the offender blames themselves and confesses to a desire to repent. By punishing the offender we are merely helping them atone. Coercing remorse and repentance may be one way in which the modern criminal system feeds itself, permitting more punishment for more crimes, while cloaked in terms of mercy.

We can add then to the individual harm suffered by criminal offenders a larger societal harm that the practice inflicts. Coercing criminal offenders into expressing support the justness of the dominant paradigm, robs society of valuable dissent. Crime displays a radical departure from society and its values, often at great detriment to the offender. If criminal offenders were not likely to suffer additional punishment for their lack of remorse then judges, juries, lawyers, police officers and society at large might hear powerful accounts of the circumstances, social failings, difficult choices, and such inherent in every crime. There is every reason to believe that such statements might lead to better judgments by the individual actors involved, and a more justice system of criminal law overall.

//

F. Conclusion

Prisoners' rights advocates justifiably seek to combat the seemingly ever growing institutions of punishment in the modern state. Any policy that potentially reduces punishment is defended uncritically. However, permitting states to grant mercy to remorseful or repentant offenders places all criminal offenders under pressure to become remorseful and repentant. This pressure is bolstered by the threat of criminal punishment in a manner that is theoretically indistinguishable from coercion. Coercion of criminal offenders into adopting a moral stance in opposition to their own character is an assault on their basic human dignity, a harm that is distinct from punishment itself. It may also have the unintended effect of increasing the palatability of punishment overall, even as it ostensibly decreases it through one act of the state's mercy at a time.

CHAPTER III: THE RIGHT TO BE IMMORAL (OR AT LEAST THINK IMMORAL THOUGHTS)

> *Is your wisdom such as not to realize that your country is to be honored more than your mother, your father, and all your ancestors, that it is more to be revered and more sacred, and that it counts for more among the gods and sensible men, that you worship it, yield to it, and placate its anger more than your fathers? You must either persuade it or obey its orders, and endure in silence whatever it instructs you to endure, whether blows or bonds, and if it leads you into war to be wounded or killed, you must obey. To do so is right, and one must not give way or retreat or leave one's post, but both in war and in courts and everywhere else, one must obey the commands of one's city and country, or persuade it as to the nature of justice.*
>
> -Plato, *Crito*

In Chapter II I stated that there is a sense in which all crime is political; that all criminal statutes are the expression of political power; and that all criminals are in rebellion against the sovereign. In this chapter I try to make good on that claim by approaching the issue from a variety of directions.

In Section A, I explain what it means to say that the criminal law is political. This does not mean that there are not moral dimensions to crime in addition to political dimensions. Rather it means that the criminal law is not an idealized reflection of morality; for good reason. Instead, the criminal law is better understood as inherently reflecting existing power dynamics in addition to morality.

In Section B, I explore one account of ideal vs. non-ideal theories of justice to show how crimes that on their face seem wholly disconnected from politics might still manifest a political dimension. Tommie Shelby has argued in a provocative book that where the existing power dynamics of a given society are sufficiently unjust, then crime becomes a permissible, perhaps even obligatory, manifestation of the principles

of self-respect and mutual aid.[33] While Shelby suggests there are clear limits to the types of crimes that fit his definition, I press on his stopping point.

In Section C, I turn to a more general theoretical discussion of the law and the way that interpretation might imbue every legal judgment with a political dimension. I argue that modern liberal constitutional theory both underappreciates this interpretive aspect of the law generally and fails to acknowledge the purpose of existing practices to address this fact. Every legal judgment is dependent on interpretation. At some point specific, abstract rules no longer obviously govern what one is to do in an individual case. At that point, the sovereign authority exerts itself in making an interpretation. This contribution is unmistakably political and reflective of the existing power dynamics in the society.

In Section D, I explore Jeremy Waldron's attempt to establish what it means to say there is a right to be immoral. Waldron's work focuses on the logical and linguistic properties of that sort of proposition. I argue that Waldron's approach is less persuasive than one that prioritizes the right to have guilty thoughts, especially thoughts that are political beliefs or attitudes.

In Section E, I link all of these approaches back to the problems of remorse in the criminal law that I raised in Chapter II.

//

//

[33] Shelby 2016.

A. What does it mean to say that all crime is political?

In stating all crime is political I do not mean to make any claims about the motivations or mental states of criminal offenders. I do not mean that every offender is motivated to communicate a political view by their crime. Nor do I mean that every offender has even thought through the social contexts of their crime and reasoned that it is in some sense justified. There are cases where some have done so, hoping that their crime will garner attention for their views and possibly bring about some sought after change. But this is likely a small subset of criminal offenses. Very likely most crimes are committed with more basic motivations (e.g. anger, desperation, panic) and/or different calculations (e.g. indifference to the law, revenge, opportunism, etc.)

I also do not claim to be making any claim about the motivations of the state generally or any specific actor in the criminal justice format. I am not claiming, for instance, that every criminal is punished because the state views them as a threat to the sovereignty of the state, such as when dictators imprison critics and opponents who are not guilty of any real crime. I am not trying to identify what are typically called political prisoners or political terrorists and claim that every offender fits those descriptions in some way.

One of the things I do mean to say in claiming all crime is political is that crime and the criminal law are distinguishable from general considerations of morality. But here I must pause and state what I am not claiming once again. For it might be said that the primary concern of philosophy of law has been the nature of the

relationship between law and morality. Two major schools have emerged: 1) the natural law tradition claiming that legal authority derives from law's connection to morality (usually derived through reason); and 2) legal positivism which denies natural law's claim and, to the extent it does anything more, claims that legal authority rests on empirical facts about human behavior and attitudes. Though I will be distinguishing law and morality in this section of the paper, I do not take myself to be weighing in on this deeper debate about the law generally. I do not intend to make claims about the ultimate source of legal authority. Instead, my main concern is to describe the way in which one subsection of the law, the criminal law, cannot be understood to be synonymous with morality. Criminal law and its enforcement judgments are not an idealized reflection of morality. Accordingly, the actions of criminal offenders should not be thought of only in terms of their morality. There is a political dimension to criminal offenses.

For instance, criminal codes do not punish all actions that a society deems immoral. Consider lying. Most consider it immoral to lie in many contexts, but the law only punishes lying in very limited contexts (e.g. lying to illegitimately gain another's property in fraud or lying about one's qualifications in impersonating a doctor or police officer). It's tempting to think that the criminal law is tracking morality in this distinction. Perhaps it is distinguishing here between lies that can cause a lot of harm and lies that might not in a way that tracks various forms of utilitarian morality. Or maybe the law is tracking a sort of Kantian distinction of lies from permissible evasions that do not harm. One example being the sort of

misleading truth Kant told King Wilhelm II when ordered by the king to cease publishing pronouncements on the subject of Christianity, Kant replied "As your Majesty's faithful subject, I shall in the future completely desist from all public lectures or papers concerning religion" knowing that the king was old and infirm and Kant would be free to resume in a short while when he was no longer the king's subject upon the king's death.[34] But this clearly is not the case as the criminal law has no interest in a great many obvious lies that cause exceeding amounts of harm (e.g. lying about infidelity, which often cause both great emotional harm and sometimes financial harm when it leads to divorce). There are many types of actions that seem immoral and that the criminal law does not punish.[35]

Additionally, the law does punish conduct that is not obviously immoral. One might think of certain traffic violations. Suppose you are in one of the most remote parts of the Mojave National Preserve between Joshua Tree and Las Vegas and you use a drone to survey and ensure there are no other people or cars within fifty miles before steering your vehicle on the wrong side of the road for a few miles. It would take a very stringent rule utilitarian or a true fanatic about civic obligations to describe this action as legitimately immoral as opposed to permissibly criminalized. And yet it is a crime and would be punishable if you turned in your drone's footage and confessed. The criminal law is not co-extensive with morality. It permits some

[34] Sandel 2009: 135-138

[35] See the examples from Waldron 1981 set out in Section D below.

type of immoral actions while justifiably punishing other actions that appear a lest morally neutral/permissible.

Here it is also important to stop again and clarify. There is a history in the philosophy of law of debating whether it is permissible for the criminal law to justify any of its provisions by reference to moral law, whether the criminal law only accidentally overlaps with areas of the moral law in pursuing other ends like preservation of order and convenience, and/or whether all questions of morality should remain outside of the criminal law in the private sphere of individuals. These questions are, perhaps, most famously taken up by Lord Patrick Devlin and H.L.A. Hart in debating English indecency laws.[36] In drawing attention to the political dimensions of the criminal law I am not arguing that the criminal law does not, justifiably, have moral dimensions as well. I agree with Lord Devlin and take it as an empirical fact that the criminal law has moral content and that it necessarily reflects a society's moral judgments, though often imperfectly. And also that a cogent theoretical account of punishment and justice in the context of the criminal law ought to have moral content. But, of course, I also agree with Hart that the moral content of the law must be subject to critical reasoning and limitations; and I abhor Devlin's specific account of morality which justifies the imposition of Christian morality on non-Christians and argues that things like homosexuality can be justifiably criminalized. All of that said, to say that all crime is political does not mean that

[36] Devlin 1965 and Hart 2014 (first published 1959).

there are not moral dimensions to criminal statutes, judgments, justifications, and actions. Any criminal offense can be considered from the perspective of morality. To say that there is also always a political perspective available is not to deny that.

It is not just the content of the criminal codes that are incongruent to an extent with morality, but also enforcement practices. Morality would require that offenders who commit similar offenses receive similar punishments and that enforcement agencies make equitable efforts at enforcement across geographical areas, populations, and economic classes. Decades of social science have proven that this clearly is not the case. Enforcement practices usually have much more to do with demographics including race and wealth, arbitrary geographical boundaries, and so on such that enforcement is undeniably uneven/inequitable to an extent that is tests what is morally permissible.

These uneven enforcement mechanisms then reverberate further such that future enforcement actions are further tinged with politics rather than morality. "The cycling of people from ghetto to prison and back again spreads a criminal ethos, an outlaw subculture, throughout many poor urban areas." In turn, "[t]he norms that govern the world of street crime also have an enormous impact on ghetto residents who want to avoid participating in and being the victims of crime. For example, the widespread use of guns by drug dealers and muggers creates a demand for these weapons in ghetto neighborhoods. Many residents, including children, arm themselves for protection, believing that the police cannot be relied upon to provide

adequate security."[37] Possession of such guns are crimes themselves in some instances, but also grounds for enhanced punishment when present during the perpetration of other crimes.[38]

What accounts for the incongruity between morality and the criminal law should broadly be understood as politics. The criminal law and its enforcement are not mere theoretical considerations. They are a series of coordinated collective actions historically situated. Accordingly, they are influenced by all of the contingencies that affect any human endeavor: competition for resources and influence, popular movements and shifts in paradigm, accident, and so on. To put it another way, the criminal law is a reflection of human power dynamics both in the long run of thousands of years of war, conquest, famine, pandemic, etc.; and in the short term where in liberal democracies legislators are elected to draft criminal laws, executives elected to fund enforcement, district attorneys and police chiefs (and often judges) elected to interpret laws and prioritize enforcement and so on.

For some criminal acts or practices in the legal system this reflection of power dynamics is easy to notice. Prohibitions on domestic violence (along with pressure on the relevant political actors to actually enforce them) and hate crime enhancements are relatively recent developments that coincide with increased political participation of historically marginalized groups. In less congenial accounts, the over criminalization and draconian punishments condemned in our justice system

[37] Shelby 2016: 206-207.

[38] *E.g.* 18 U.S.C. § 922(g)(1) and *California Penal Code* § 120022

today can be traced to anxieties of white elites following the emancipation of African American slaves (and the current pushback might also be seen as reflecting another reactionary shift in power).[39] We are able to easily see the political dimensions in these cases because the shifts in the underlying power relations are recent and obvious.

At this point, a cynic might remark that claiming that criminal law is incongruent with pure morality and tainted by politics isn't a very illuminating claim. Rather it is common knowledge and evidence of human imperfection. There are two responses to such a comment. The first is that while it may be common knowledge, it is rarely acknowledged in philosophical accounts of the criminal law. Instead, commentators on the criminal law tend to focus on abstract, ahistorical moral considerations to delineate what is wrong or right about specific criminal acts and procedures. Doing so results in anemic accounts, which I think is revealed by Shelby's work discussed in Section B below. Second, the cynic's view portrays the political dimension of crime as a bug and not a feature of the criminal law. The view seems to be that the aim of criminal law is simply to enshrine rational morality. What this view ignores is that the application of any code to unique, individual cases requires interpretation; and where the authority resides to determine edge cases will always be a political question. This is the topic of Section C below.

As it appears that modern liberal states do not punish all acts that are deemed immoral by that society, and also punish some acts that are not considered immoral

[39] Alexander 2012.

outside of their specific political context, then it seems clear that criminal laws are not merely the formal recording of a community's particular view of morality. The most plausible view of what accounts for each of these differences is that the existing power dynamics of a society play an important part in its criminal justice system. Politics determine, at least in part, both what actions are deemed criminal and, equally important, how resources will be allocated for enforcement. To claim that all crime is political is also to claim that this broader context must be considered when evaluating the actions of offenders and/or criminal justice practices. Provocative work on ideal vs. non-ideal theory illustrates this point.

B. Ideal v. Non-Ideal Theory

Shelby views himself as a committed liberal-egalitarian theorist in the tradition of Kant, John Stuart Mill, and John Rawls. He goes so far as to claim that to his mind ideal liberal-egalitarian theory has been given a thoroughly detailed and convincing defense. Shelby views this tradition as having legitimately set the "practical goals to work toward and as normative standards for judging the overall justice of particular social arrangements."[40] Of course, the normative standards of the tradition Shelby reveres are far from actualized, especially with respect to residents of economically disadvantaged and racially segregated ghettoes. Shelby's project is therefore one of non-ideal theory that respects the liberal-egalitarian ideal tradition, but is responsive and useful to the actual existing historical context.

[40] Shelby 2016: 12.

Non-ideal theory "specifies and justifies the principles that should guide our responses to injustices."[41] (Shelby 2016: 11) The tenets of ideal theory serve as standards and/or aspirational goals, while non-ideal theory governs how states and individuals are to handle instances of failure or structural patterns of failing to meet those standards. Thus non-ideal theory is relevant to the criminal law in both proscribing where and how the state may punish and individual responsibility to obey the law. How far the context strays from ideal justice will matter.

Shelby follows Rawls in locating the duty to obey the law in both civic obligation and natural duties. "Civic obligations are owed to those with whom one is cooperating in order to maintain a fair basic structure. They are the obligations that exist between citizens of a democratic polity as defined by the principles of justice that underpin their association. Civic obligations have binding normative force because of the contingent associational ties between citizens, that is, because of the formal or informal bonds that define a set of persons as a distinct people or nation. By contrast, natural duties are unconditionally binding, in that they hold between all persons regardless of whether they are fellow citizens or are bound by other institutional ties. Both civic obligations and natural duties are moral requirements. The key difference is that one has civic obligations qua citizen and natural duties qua moral person."[42]

[41] Shelby 2016: 11.

[42] Shelby 2016: 213.

What emerges from Shelby are descriptions of the injustices inflicted upon poor, largely African American ghetto residents and the ways that they might relieve them of their civic obligations, and possibly even obligate them to express their resistance through certain types of crime. "Unjust social arrangements are themselves a kind of extortion, even violence, and consent to them does not bind."[43] To this axiom of ideal theory Shelby appends two non-ideal principles of resistance to injustice: 1) the ethical duty of self-respect; 2) the ethical duty of mutual aid. Self-respect requires recognizing and affirming one's equal moral worth as a person. "[M]y articulation of an ethics of the oppressed is perfectly consistent with what might be called the *Stand and Fight* tradition in black political thought. This tradition counsels against suffering in silence and insists on fighting openly and assertively—in the press, in the legislative halls—in the courts, and in the streets—and not only quietly behind the scenes."[44] The principle of mutual aid is the duty to help the needy, vulnerable, and weak when one is able.

Of course, it doesn't mean all criminals have dissent as their primary conscious motivation. "People from all races, classes, and types of neighborhood engage in criminal activity for money, status, power, or amusement. When poor persons from ghettos choose crime, however, they do so under conditions of material deprivation and institutional racism. Thus their criminal activity sometimes expresses something more, or something other than a character flaw or disregard for the

[43] Rawls 1999: 220.

[44] Shelby 2016: 6.

authority of morality."[45] That something more is a political belief about injustice and crime as a means of dissent.

Shelby is very careful to extend this analysis to only certain types of crimes for while "serious, burdensome, and repeated injustices over the course of a citizen's life can release them from these civic obligations, it cannot from their natural duties."[46] While unlawfully taking the possessions of another might be permissible, natural duties likely still prohibit the use of violence to do so in almost all cases and the ethic of mutual aid requires that one not take advantage of people victimized by the same systemic injustice as the offender. "Something similar can be said in favor of prostitution, welfare fraud, tax evasion, selling stolen goods, and other off-the-books transactions in the underground economy."[47] Mutual aid may render participation in group criminal activity like gangs morally permissible, or even required, but not the active recruitment of or predation on the vulnerable.

The terms are all Rawlsian, but there is more than a whiff of Hobbes in Shelby's approach. Shelby is concerned with understanding and improving the circumstances and obligations of ghetto occupants in modern America. Social scientists relied on by Shelby have done significant work cataloguing the many severe types of injustice unfairly imposed, willfully or negligently, on this specific group.[48]

---

[45] Shelby 2016: 205.

[46] Shelby 2016: 214.

[47] Shelby 2016: 220.

[48] Alexander 2012.

However, it does not take much imagination (or more than a passing familiarity with world history and current events) to imagine social contexts that are even more unjust than the milieu Shelby focuses on. In social contexts that include state sanctioned slavery or genocide principles of self-respect and mutual aid likely justify much more extreme crimes than the ones Shelby is willing to consider possibly justified for the ghetto poor. More extreme crimes are even consistent with the idea of natural duties in some renderings as they are justified under a principle of resistance very similar to, or even derivative of, Shelby's principles of self-respect and mutual aid: self-defense and defense of others. That belief animates our widely held (and legally codified to various extents in every part of the United States) intuition that self-defense and defense of others justify the use of deadly force in some circumstances.[49] Where the state has inflicted social injustices on a segment of the population so severe that they approximate a state of war, then members of that population must surely be authorized to take up arms in their own defense, to sabotage the state, and to prevent its agents from advancing its policies and programs (at least those that are unjust) by any means necessary. Now we have reached a point where the crime of murder, which is perhaps most resistant to the claim of being political, has to be understood within its specific political context. For there will always be a question of whether the crime is objectively reasonable given the conditions of the society. And while a

[49] *cf.* People v. Goetz, 68 N.Y.2d 96, NY Ct. of Appeals (1986); *Model Penal Code* § 3.04; and United States v. Peterson, 483 F.2d 1222, Ct. of Appeals, District of Columbia (1973)

majority may find it easy to answer "no" in any specific case, it is important to keep in mind that that is a judgment that has a political aspect to it, and respect for the criminal offender's rationality requires us to view their action as the expression, perhaps imperfectly, of a dissenting political viewpoint.

Thus, the more unjust the criminal justice system and the total social conditions in a society, the more reason we have to believe that crime itself is partially a political act and not merely a moral failing of character. And of course this argument applies not just to the content of the law, but to the way the law is applied and enforced; the motivations, attitudes and biases (conscious and unconscious) of the various actors that participate in it, and the consequences that enforcement has on the individuals, subgroups, and society overall. But even if a system of criminal justice and its social context were to approximate justice to a significant degree we might still have reason to think that all crime is political to some extent. The reason lies in the interpretive nature of law.

C. Interpretation and Sovereignty

The nature of interpretation in legal reasoning, the proper methods of interpretation to use in legal contexts, and the threat, if any, that interpretation poses to legal constitutionalism are some of the most active areas of research and debate in the philosophy of law over the last forty years. Engagement with all of this work is not possible here, nor is it necessary to the minimal observation that I wish to make: that judgments of guilt in the criminal law are political judgments informed, at least in part, by political beliefs.

There are four types of legal judgment relevant to the criminal law: 1) determination of the content of the law; 2) determination of the facts at issue in a particular case; 3) judgment of how to apply 1 to 2; and 4) judgment about how a case should be decided considering anything and everything else. In separating these types here I do not intend to weigh in on the debate as to whether there are actually distinct bifurcated phases/stages in legal reasoning and/or whether there is any merit to the contention that legal judgment consists of determining whether any rules apply and then effectively legislating in the spots where they run out. Nor am I interested in whether there is a general theory of legal interpretation that can be used to guide decision makers in judgment types 1-3 or whether there is one right answer in those cases.[50] I am only concerned with drawing attention to judgments of type 4 here, and how they are unique because they permit the adjudication of a case on basically any grounds the decision maker deems appropriate. And they are made in every criminal jury trial in the Anglo-American tradition. The wide open nature of type 4 judgments means that in every conviction or acquittal by jury there is an implicit political judgment rendered.

Anglo-American courts have long recognized the prerogative of the jury to disregard evidence, any instructions about the content of the law they receive from the judge or lawyers, and to acquit a criminal offender. This is called jury nullification and is essentially the unfettered right to make type 4 judgments in

[50] *Cf.* Dworkin 1977 & 1986; Raz 1996a & 1996b for the accounts of two of the most prominent commentators on legal interpretation on these foundational questions.

acquittals. Jurors are not required to explain their reasoning in reaching these type 4 judgments and acquittals are not reversible by the court. Jurors are not subject to any penalty for outrageous decisions absent bribery or some other form of corruption.[51] The judgment to convict, on the other hand, is restrained by the burden of proof beyond a reasonable doubt that lies with the state. However, the judgment to convict can be understood to be a rejection of the sort of political beliefs that might lead to an acquittal and therefore it contains political content in the same way.

One way to think about jury nullification is that it is the embodied right to make the sort of calculations that Shelby urges us to make as we evaluate the actions of criminal offenders. For instance, jurors are permitted to find that a criminal offender did in fact embezzle from their employer in a way that is an offense under the law, and yet acquit because they also find (or merely know from their own experience of the society) that it is unfair for the state to criminalize or punish this behavior since there are unjustly limited job opportunities for the offender, unjustly dehumanizing working conditions in this specific employment, the money taken was insignificant to the stakeholders in the business yet very significant to the offender, and so on. These calculations are informed by political beliefs and attitudes about questions like what levels of economic inequality are permissible in a just society, what are the obligations of society to correct the historical injustices of slavery and segregation that remain enshrined in current policies and practices, when is

[51] See United States v. Dougherty, 473 F.2d 1113 Ct. of Appeals, District of Columbia (1972) for a concise history and summary of jury nullification.

enforcement of criminal penalties so unevenly applied that it can no longer be justly applied in an individual case.

I have made the point so far with respect to the judgments of juries because the existing law in America is very clear that juries have the authority to engage in type 4 reasoning. But it is also present in the discretion of sheriffs to refuse to enforce certain crimes, of prosecutors to refuse to charge and try certain cases, and of the discretion judges have to permit or exclude evidence in trials on the grounds that it is unduly prejudicial. At every step of the path leading to a criminal conviction the political beliefs of the various actors involved have informed the process.

A possible objections arises here. First, someone might claim that all I have done is reveal the deficiencies of modern American criminal procedure. Efforts should be made to exclude the political beliefs of actors in the criminal justice system. There are two responses. The first is just to note that the criminal procedure as I have described it has existed for hundreds of years and, despite complaints over that entire period, it does not appear to be changing anytime soon. Therefore, it is valuable to note this feature as we evaluate other aspects of the criminal law.

A more foundational response to the objection is to wonder whether it is even possible to exclude political beliefs from informing legal judgments. Foundational issues related to political interpretation include whether and where to locate political sovereignty. Liberal-constitutionalism generally assumes that rules for applying general legal rules to concrete cases can be emobodied in the material content of the law itself. "The law can determine, for any material legal norm, which person or

institution has the competence to interpret and apply it."[52] There is a plausible alternative, however, that argues that there is always the possibility of an emergency where various competing provisions in the law make this determination of competence impossible. Think of the disastrous constitutional crisis of contested election results where all three branches of government answer the question of who won differently. The only way to resolve the dispute would be to step outside the law, what theorist Carl Schmitt called an exception. As Schmitt famously[53] claimed "Sovereign is he who decides on the exception."[54] One way to understand the fact that the modern American judicial system permits type 4 judgments by elected officials (and quasi elected officials) and members of the public impaneled on a jury is as an attempt to locate sovereignty in criminal matters as closely to the general will of the citizens as practically possible. Unlike philosophical controversies, criminal court cases actually have to reach a resolution and taking polls of the entire community is impractical. Returning to the objection, it may be impossible to altogether remove type 4 judgments from criminal. We may only to relocate them to other institutions, or individuals as autocracies do, but final say still has to reside in some person or institution other than the law itself. Exceptions inevitably arise that the content of the law has not (perhaps even cannot)[55] anticipate. Type 4 judgments

[52] Vinx 2019.

[53] Or infamously, as Schmitt would go on to become a prominent Nazi whose theories would justify dictatorship. Even so, his challenge to liberal constitutionalism still engages theorists. *See* Caldwell 2005 and Vinx 2019.

[54] Schmitt 2006.

[55] And I as I argue in Chapter V should not where redefining terms like "reasonable"

are the only way to resolve such exceptions. Relocation of the authority to make type 4 judgments to juries and/or elected officials is one way to make them more democratic and just, but it does not eliminate them from criminal procedure.

So far, I have argued that the content of criminal statutes and enforcement practices significantly reflect existing power dynamics within a society such that evaluating individual criminal offenses and state practices on purely moral terms is untenable. Non-ideal theory imposes a duty to consider the actual historical context that criminal offenders are enmeshed in. Judgments in the criminal law themselves are political acts informed by political beliefs. From this standpoint, engaging in crime may be an expression of a dissenting political opinion (though how salient or developed that opinion is will inevitably vary). From this framework emerges a distinctively political right to act, in some sense, immorally or at least think immoral thoughts.

D. <u>The Right to Be Immoral and/or Think Immoral Thoughts</u>

Other thinkers have noted a similar distinction to the one that I have drawn between crime and morality. Jeremy Waldron has noticed that we often say that A is morally wrong, but that O has a moral right to do A. There is intuitive appeal in those sorts of statements, but they also have the appearance of a paradox or at least inconsistency. Waldron gives a number of examples of actions where he thinks we would plausibly want to make these statements:

---

in criminal statutes distances them from the community's ordinary understanding of the terms.

> Someone uses all the money that he has won fairly in a lotter to buy racehorses, and champagne and refuses to donate any of it to a desperately deserving charity;
>
> An individual joins or supports an organization that he knows has racist leanings, such as the National Front in the United Kingdom; he canvasses support for it among a credulous electorate, and he exercises his own vote in its favor;
>
> Somebody offers deliberately confusing, though not untrue, information about the policies of a political party to a confused and simpleminded voter in an attempt to influence his vote;
>
> An athlete takes part in sports competition with the representatives of an oppressive or racist state, despite the fact that this profoundly demoralizes those who are struggling for the liberalization of that state;
>
> Antiwar activists organize a rowdy demonstration near a cenotaph service on Remembrance Day;
>
> A man refuses to give a stranger in the street the time of day when he asks for it or coldly and rudely rebuffs attempts at conversation in a railway compartment;
>
> Someone refuses to consider evidence that might call in question his or her fundamental opinions and beliefs about the world; for instance, a biblical fundamentalist refuses even to look at the evidence of the fossil record.[56]

Waldron argues that there is no paradox in making the sort of statement he has identified in these cases because claiming that an action is morally wrong does not entail claiming that third parties have a right to interfere with O's performance of that action. And the considerations that go into making those two different determinations are distinct.

[56] Waldron 1981: 21-22.

Waldron's focus on moral rights to the exclusion of legal rights and his logical/linguistic approach to the appearance of paradox fail to illuminate the real source of intuitive appeal in the examples he surfaces. For instance, he fails to acknowledge that every single example he produces where it is plausible to say that an action is morally wrong but permissible is an instance where the offender holds a specifically political attitude; that is a belief about justice. The dilettante lottery winner clearly has a strong commitment to individual property rights; the rude railway passenger clearly has a very narrow view of the obligations owed between fellow citizens; and all of the others are expressing explicitly political views about how society ought to function, what the state ought to be like, and so on. In many of the cases, the person is not only holding a political opinion but taking non-immoral actions to foster the growth of that opinion.

To the extent that we want to say that it is not okay to interfere with performing those actions, we are only saying that we have no right to physically interfere with the political beliefs of others. For we would never say that it is morally wrong to murder and also that an offender has a right to murder and third parties have no right to interfere to stop the murderer. The only sorts of cases, at least as far as my imagination takes me, where Waldron's statement is intuitively appealing are cases where there is a political attitude at center combined with actions that are not prima facie immoral. This second point is equally important, because we would at the very least hesitate to say that punching people in the face is immoral but O has a right to punch people in the face who disagree with his political views.

While we may be permitted, even morally obligated, to interfere with certain immoral actions we are not permitted to, even morally obliged to refrain from, physically interfering with these political attitudes. I will try to further this point by exploring the criminal law's requirement of voluntary action as a precondition for punishment. That is to say, the fact that we do not punish offenders on the basis of criminal thoughts alone.

A great many criminal law commentators have opined on the reasons for why it is so, and/or just, that the law not punish for mere thoughts. With some variations in detail, they seem to fall into four broad categories: 1) we cannot reliably detect guilty thoughts[57]; 2) everyone has them and so punishing them would entail expanding the criminal law to an unreasonable or undesirable extent[58]; 3) harm is required to punish and there is no harm in mere thought[59]; and 4) some combination of the three. Each of these are deficient, leaving open the possibility that either the principle of not punishing thoughts is unreasonable or that there is some unidentified reason justifying the practice that reveals something significant about the nature of criminal responsibility in general.

1 seems to be just false. Suppose that someone confesses to their therapist that they fantasize about murdering their spouse. Do we really doubt their confession as reliable evidence that they are indeed thinking of committing a heinous crime?

---

[57] *Cf.* William Blackstone's "Commentaries" reprinted in Kadish & Schulhofer 2001: 181; Dworkin & Blumenfeld 1966: 401; Goldstein 1959: 405-406.

[58] *Cf.* Williams 1961; Stephen 1833 reprinted in Kadish & Schulhofer 2001: 181.

[59] *Cf.* Mill 2002; Feinberg 1984.

Mandatory reporting requirements make it seem otherwise. But even in this case, we wouldn't punish the person unless they had taken some action in furtherance of the plan. We might investigate to see if that was the case. We might notify the spouse to keep them safe. We might offer the patient increased therapeutic interventions. But we wouldn't punish and it certainly wouldn't be because we doubted that the patient actually had the thoughts.

2 is at best a purely logical point, *reductio ad absurdum*. Accordingly 2 would not act as justification for refraining from punishing for thoughts alone. It would merely be a pragmatic guide that we should not engage in the practice until we can untangle or re-conceptualize to find the logical error or see that there is, in fact, none to begin with. At worst it is empirically false. Is it really true that everyone has graphic rape fantasies or some other immoral thought the law might seek to punish? There is good reason to doubt the supposed explosion of the criminal law.

3 might be false and it relies on the harm principle as justification for punishment, which may also be false. There are emerging philosophical arguments that thoughts alone (such as believing something negative about a person due to their race) can be a sort of wrong we do to another.[60] There is some plausibility to these arguments, though they are inchoate. Perhaps more damning of accounts in this category is that the harm principle does not adequately justify all that we legitimately believe is just to punish. For instance, we generally do not allow victims to consent

[60] Basu 2019.

to significant violence even where that violence cannot easily be understood to harm the victim or frustrate their interests.[61] If the harm principle is not a restriction on criminal punishment, than 3 does not prohibit the punishment for mere thoughts.

And if 1-3 are deficient, I see no reason why combining them would accomplish anything more than act as pragmatic guide which shows us it would be safer not to punish for mere thoughts. It would not be a persuasive argument that punishing for mere thoughts is unjust. There seems to be some other justification for the practice, or widespread intuitions that it is wrong to punish mere thoughts are indefensible.

The passage from Crito that opens this chapter suggests an alternative. There may be a political right to think immoral thoughts, and even to act on them to the extent the action is a peaceful effort to change the status quo of the state and/or persuade one's fellow citizens to change their minds about what is just.

Consider Socrates' attitude toward his impending execution. Socrates is ready to endure the harsh treatment that is his punishment as it is the lawful judgment of the state. He will not use force or trickery to try to avoid this harsh treatment. He has earned it insofar as the state has determined he has earned it. But that does not mean that Socrates has changed his mind about what whether his actions, deemed criminal offenses by the state, were actually just. He gave an impassioned defense of his actions at his trial, he continues to call the state's judgment of his actions "faulty"[62]

---

[61] New Jersey v. Brown, 143 N.J. Super. 571 (1976).

[62] Plato 2002: 53.

while in confinement awaiting his punishment, and he tells Crito that he would happily accept not being executed so long as it was through the persuasion of the state and a lawful order acquitting him. Socrates, or Plato through Socrates, seems to hold that while it is just for the state to impose punishment on the offender for criminal actions, the offender does not forfeit their right to their political beliefs about justice. Furthermore, they are permitted to act on those beliefs to persuade their fellow citizens and state actors to change their beliefs and actions. This is the right to be immoral, or at the very least, have immoral thoughts

E. Remorse

In Chapter II I argued that coercing a criminal offender into expressions of sincere or performative remorse interfered with their individual moral right to autonomy. In this chapter I have tried to articulate a related political right to thoughts that the state might deem immoral and to some action in furtherance of those beliefs. It requires allowing criminal offenders to continue to hold and peacefully express whatever political beliefs or attitudes that they choose to without forceful intervention. These beliefs and attitudes are not an excuse that absolves them of blame or punishment for criminal offenses or immoral actions, but it is also not a legitimate use of political power to attempt to violently change these opinions or stifle their participation in public discourse, such as at a sentencing hearing.

To say that all crime is political and all prisoners are political prisoners is to acknowledge the many ways that any system of criminal justice reflects existing

power dynamics that are often difficult to see, but can drastically affect our calculations as to whether blame or punishment is warranted at all.

In the same way that it is wrong to strip criminal offenders of their right to vote, it is wrong to use violence to coerce them into sincere or performative expressions of remorse. They should be respected as political participants and provided with an opportunity at sentencing hearings to honestly explain the motivations and calculations that led to their criminal offenses without fear of incurring punishment beyond which they might receive if they apologize and condemn themselves. It just might be illuminating political discourse that leads to a more just society.

"Just as physicians take basic human anatomy as given when treating patients, policy makers working within the medical model treat the background structure of society as given and focus only on alleviating the burdens of the disadvantaged. When it comes to the ghetto poor, this generally means attempting to integrate them into an existing social system rather than viewing their unwillingness to fully cooperate as a sign that the system itself needs fundamental reform. In short, features of society that could and should be altered often get little scrutiny…In addition, the technocratic reasoning of the medical model marginalizes the political agency of those it aims to help. The ghetto poor are regarded as passive victims in need of

assistance rather than as potential allies in what should be a collective effort to secure justice for all."[63]

Returning to the three types of criminal offenders I have distinguished we can understand our patterns of response to them in a new light. Political prisoners are offenders who have made their political beliefs salient in the commission of their offense, and whose criminal offense is largely peaceful (such as disrupting traffic temporarily or occupying other public spaces peacefully), and likely not more than a breach of civic duty which might be justified given the context. We recognize that their actions are not so egregious that they demand a state response. Invariably, our sympathies are strongest where their political beliefs are consonant with our own, but even when they're not we instinctually recognize that the state must constrain itself lest it illegitimately persecute citizens more for their beliefs than their actions.

Political terrorists are offenders who have made their political beliefs salient in the commission of their offense, but the content of those beliefs does not matter to us because we feel constrained to blame and punish them for offenses that are so heinous and unjustified in all but the most extreme of contexts, or possibly never at all. To clarify, the content of their political beliefs will not provide them with an excuse whether we agree with their beliefs or not. We recognize we are punishing them for their action and not their beliefs. We instinctually recognize that it would be wrong to impose force merely to change those beliefs.

[63] Shelby 2016: 2.

For the remaining criminal offenders, which surely make up the vast majority, we willfully ignore the political dimensions of their crimes. Worse, we threaten them with violence if they choose to make them salient at some point after the commission of their crime. By coercing remorse we forcefully mask the political nature of their actions and possible motivations. This is a violation of the right to think immoral thoughts and to participate peacefully in public discourse. It is a breach of the moral and political rights of offenders and it prevents the criminal legal system from achieving its end of justice.

CHAPTER IV: A GENEALOGY OF REMORSE IN THE CRIMINAL LAW

> *Duke: Thou shalt see the difference of our spirit; I pardon thee thy life before thou ask it: For half thy wealth, it is Antonio's, The other half comes to the general state, which humbleness may drive unto a fine…*
>
> *Shylock: Nay, take my life and all, pardon not that—You take my house, when you do take the prop that doth sustain my house: you take my life when you take the means whereby I live.*
>
> -William Shakespeare *The Merchant of Venice*

Why does the state consider a criminal offender's expressions of remorse, either sincere or performative, in determining how much punishment to inflict upon the offender? After all, remorse is not obviously connected to desert or blameworthiness. Any retributivist who wishes to justify the practice is forced into ungainly contortions that put at issue the offender's entire character rather than their blameworthiness for a specific crime. Retributive accounts that would accommodate remorse struggle with proportionality. Punishment either cannot be imposed at all where an offender is sincerely remorseful or imposed to an absurd extent in order to reform an offender's moral character. Utilitarians might claim that less punishment is warranted because the expressions produce some good. Maybe remorseful offenders are less likely to reoffend or that their expressions offers some benefit to victims. However, there is a lack of empirical support for utilitarian ends that would justify the practice.[64] Utilitarian theories also struggle to account for the way the practice

[64] *See* Bagaric &Amarasekara 2001 for a summary of empirical studies that could not find a reliable connection between remorse and reduced rates of recidivism. The authors also argue that remorse should not be considered in criminal punishment

operates in real cases. For instance, why do courts consider remorse in cases where there is no obvious victim or concerned relative, or when the victim is unaware of the crime.[65]

Another way to phrase my initial question in this chapter: if the practice is not obviously supported by either of the two major theories that justify punishment, then how did the practice become entrenched and what is the source of its intuitive appeal? For it seems to have snuck into the castle unnoticed by the criminal law's theoretical gatekeepers. Another possibility: it made its way inside before the guards were even posted; before even the need for guards was realized. And it has been posing as one of the legitimate citizens of the kingdom undiscovered ever since.

In this chapter I articulate this possibility through a philosophical genealogy of remorse and punishment. In Section A, I examine remorse as it may have progressed from interpersonal relationships into the criminal law. In Section B, I argue that the intuitive appeal of remorse in the context of the criminal law may lie in retributive emotions. This should give us pause, and I consider whether such emotions can give rise to virtuous judgments. In Section C, I consider whether the

---

because the practice is not supported by any of the major theories for the justification of punishment. While I agree with their basic point, much of their thinking is inconsistent with my own. Their work lacks empathy for criminal offenders generally. They display little to no concern for the moral or political rights of prisoners.

[65] Consider a case where the state punishes an offender for possession of child pornography, but the victim of the sexual assault lives in a faraway country and has no knowledge of offender or their actions.

principles articulated in Sections A & B are compatible with the demands of what is sometimes called restorative justice.

A. A Genealogy of Remorse

I intend the account of remorse I advance in this section in the same spirit of philosophical genealogy that Rousseau applied to property[66], Nietzsche applied to morals[67], Charles W. Mills applied to a racial contract which shadows the social contract,[68] and, most importantly, Foucault applied to the modern prison industrial system.[69] These genealogies are both descriptive and normative. They are descriptive in that they do provide at least sketches of supposed factual/historical evolutions, and are subject to being evaluated as accurate or inaccurate. However, their value lies primarily in their normative power. As Mills puts it they are "normative tool[s], conceptual device[s] to elicit our intuitions about justice." These sorts of genealogies then can be criticized along both of these dimensions, but undermining the value of one will not always undermine the value of the other. Quibbles with the accuracy of the descriptions (so long as they do not undermine the overall plausibility to an outrageous degree) may not undermine the genealogy's value as a lens for thinking about justice; and determining the genealogy is not a fruitful approach to normative judgments will not necessarily mean it is flawed as a

[66] Rousseau 2000.

[67] Nietzsche 1969.

[68] Mills 1999.

[69] Foucault 1995.

history. It is my hope then, that with the brief account I offer here any descriptive deficiencies are overshadowed by the normative value of the story I tell.

An Individual

There is a biological evolutionary story that one could tell about how among the animals on Earth developed the pattern detector *par excellence*: the human brain. And that in sorting chaos of detectable phenomena the correlation of one thing being followed by another thing repeated frequently enough for causation to emerge. From there flowed pleasure, pain, self, plan, action, success, failure, misfortune, agency and so on until you eventually get to responsibility and regret. I do not have much to contribute to that story (nor am I qualified), and so I will refrain. But it is worth noting that while responsibility and regret make sense in this primordial context of the individual it is much less certain that remorse does. For remorse is not merely regret at failing to achieve one's intended effects through one's actions. There is no sense of remorse in the feeling of throwing an object at a target and missing. As Christopher Kutz has catalogued there are many senses of responsibility and regret that do not carry the ethical judgment that is inherent in remorse.[70] For remorse to emerge there must be concepts such as role, duty, others, harm to others and so on. These are only possible in the context of our relationships with others.

//

//

//

[70] Kutz 2004.

Interpersonal Relationships

Murphy gives an account of what he takes to be the most natural context for remorse and forgiveness: interpersonal relationships.[71] In our private relationships we occasionally cause harm to others we care about. When we harm someone like that we threaten to rupture the relationship. If the aggrieved party cannot move past the harm, the relationship will end. Forgiveness is judging that you have been harmed, but agreeing to continue on with some form of relationship with the offending other. One good reason the aggrieved party might not forgive is self-respect. It is a vice to continue on as normal with someone who causes you harm and does not experience regret. Remorse is sort of an ethical precondition to forgiveness on this view, one that it would be wrong not to demand where the conduct that caused the harm was blameworthy. By demonstrating moral regret for their blameworthy action, the offending party to the relationship restores and affirms a proper respect for the other. On that basis the relationship can go on. Being able to experience remorse is a virtue as it makes one eligible for forgiveness and capable of enduring relationships.

From this basic account it's possible to make sense of the concept of self-forgiveness as well. The need to receive forgiveness from others helps to reveal the ways that we have failed to meet our ethical obligations to them. Virtuously expressing remorse makes us eligible for forgiveness from others, though not always

[71] Murphy 2003.

entitled to it. But even after we receive it we might still feel an inner guilt for our failure to meet our ethical standard. Remorse is a way of reaffirming our standard such that we can go on with ourselves less burdened. Forgiving wrongdoers may "enable [them] to forgive themselves by showing them there is still enough decency in them to warrant" it.[72]

As human beings developed the complexity of relationships, roles, duties and standards of care exploded. Returning to the many sources of responsibility articulated in Kutz, we can see that the concept of remorse as articulated above does not map onto all of the instances where humans cause harm. When an employee recklessly damages the employer's merchandise is remorse really required in order for the employment relationship to continue? Since the primary basis of the relationship is mutual financial gain instead of something more personal like mutual care/affection/love it does not appear that remorse is ethically required here. An employer is not obligated by any ethic of self-respect. They may not even be human, but a corporation or government. All that is needed for the relationship to go on is the employer to recognize the financial gain in continuing the relationship outweighs the cost of damaged goods that results from their occasionally blameworthy behavior.

The Inquisition

In the next part of the story, the virtue of remorse, along with many other virtues, becomes codified in formal religion. As religion becomes more organized it warps remorse so that the proper target of remorse is no longer those who are harmed,

[72] Murphy 2008: 70.

but the source of the church's authority, God Himself, and therefore it is the Church who has the power to adjudge guilt and grant forgiveness. Governments recognize the utility of religion in controlling citizens and import whatever values and beliefs are useful. This partnership climaxes in the Catholic Inquistion of the Middle Ages.[73] In practice, the state monarchies and the established church partnered together to eliminate threats to either of their power. The Inquisition would supply the justification for punishment and the state would supply the actual force. In particular, the Inquisition would torture victims to extract confessions, and then use those confessions to justify having tortured in the first place. The logic behind the practice was that repentance was in the spiritual interest of the offender. Remorse was transfigured from a private moral virtuous practice into a weapon of control.

The Secular Criminal Justice System

Eventually this partnership erupts into a conflict for supremacy between church and state. It results in the secular criminal justice system, but the practices of coercing remorse out of criminal offenders was already deeply entrenched. As Foucault documents, efforts at humane reform often end up serving the interests of established power structures and dynamics.[74] Torture of prisoners in the town square is a conspicuous use of the violence by the sovereign that might arouse the ire of a citizenry who disagrees with its justness. Whereas prisons are "gentler" forms of

[73] Murphy 2013.

[74] Foucault 1995.

punishment, "to punish less, perhaps; but certainly to punish better."[75] Corporal punishment is abolished and the modern prison is relocated to remote locations where the citizenry is not constantly confronted by the violence maintaining the order of the state. The judgments that permit the punishment are no longer so obviously questionable (like the tenets of religion had become). Instead they are scientific judgments of normal and abnormal. But the underlying logic remains, the violence of the state is not violence at all, it is a service to the offender. A sort of cure.

In this way a virtuous attitude of responsibility and respect in the private interpersonal context transforms into a method of control that serves existing power structures and dynamics. The intuitive appeal may come from our familiarity with remorse in its unproblematic private interpersonal context. Further intuitive appeal of the practice likely lies in our conditioning to and varying levels of interest in preserving these structures and dynamics. In the next section of the paper I examine another reason we may be drawn to the practice: the retributive emotions. Collectively holding a defenseless criminal in our power and extracting remorse may serve our more vicious emotional needs than any legitimate just interest in blameworthiness, mercy, or justice.

B. <u>The Retributive Emotions</u>

One of Nietzsche's most profound insights in moral psychology is that our intuitions about which actions and attitudes are generally believed to be virtuous may

[75] Foucault 1995: 67.

have their origins in emotions that are generally believed to be immoral. In this Section of the paper I will argue that the appeal of our ability to grant mercy to criminal offenders may merely be masking the activation of retributive emotions. The structure of this section is as follows: 1) I will summarize Nietzsche's critique of the retributive emotions in the context of the criminal law, in particular how they distort our ability to see the actual blameworthiness of a criminal offender and damage the virtue of those in judgment of the offender; 2) I will critique Michael S. Moore's response to Nietzsche in which Moore tries to rescue the retributive emotions as epistemically relevant to blameworthiness; and 3) I will try to discern whether any principles from Nietzsche suggest a way we might transcend the retributive emotions.

1. Nietzsche's Critique of the Retributive Emotions

Moore refers to the retributive emotions as dubious, murky origins for epistemic judgments: "a witches brew…that Nietzsche sometimes lumped under the French term *ressentiment.*"[76] In particular, Nietzsche identifies weakness or impotence, fear, herd morality, sadism, and general self-deception as potential origins for retributive emotions.

Weakness/Impotence

At times in his writing Nietzsche identifies the criminal with strength.

> The criminal type is the type of the strong human being under unfavorable circumstances: a strong human being made sick. He lacks the wilderness, a somehow freer and more

[76] Moore 1988:192.

> dangerous environment and form of existence, where everything that is weapons and armor in the instinct of the strong human being has its rightful place. His *virtues* are ostracized by society; the most vivid drives with which he is endowed soon grow together with the depressing affects—with suspicion, fear, and dishonor.[77]

If the criminal is strong, then the implication is clearly that those who would blame them are weak. Moreover, the criminal may not be actually blameworthy at all, but merely in circumstances that are not suited to the criminal's strength (i.e. in society). In the wild or at war the very traits that arouse fear and suspicion of the criminal would be hailed as great virtues. Thus, it is not the criminal's actual blameworthiness that gives rise to retributive emotions, but the recognition of those who blame that they are weak and afraid of the criminal.

Nietzsche goes even further in his *Genealogy*. It is not merely weakness that can give rise to retributive emotions, but impotence to act in a similar manner. Nietzsche illustrates the point with the metaphor of harmless lambs and the birds of prey who eat them.[78] The lambs come to the conclusion that they can blame (or label "evil") the birds of prey for eating them by reasoning that the lambs forbear from eating the birds of prey. Nietzsche's point is that obviously this line of reasoning is unsound and self-serving. The lambs do not forbear from eating the birds of prey, they are incapable of doing so. The source of the lambs' retributive emotions is not actual blameworthiness, but rather their experience of their own impotence in

[77] Nietzsche 1954b: 549-551.

[78] Nietzsche 1969: 44-46.

comparison to the birds. Often blame is merely "the oppressed, downtrodden, outraged exhort[ing] one another with the vengeful cunning of impotence: 'let us be different from the evil, namely good!'"[79] But of course, the lambs are not merely expressing their weakness/impotence in blaming the birds of prey, but also their fear. To hear criminal offenders express remorse strengthens this self-satisfying illusion for non-offenders, regardless of the actual blameworthiness of the offenders.

Fear

As alluded to above, when retributive emotions have their origin in weakness/impotence they also tend to derive from fear. Obviously fear is a source of retributive emotions that may not be actually related to actual blameworthiness, but as importantly for Nietzsche fear can also significantly increase the retributive emotions beyond all justification. In some instances, punishment of a lawbreaker devolves into a forum where "every kind of hostility may be vented upon him. 'Punishment' at this level of civilization is simply a copy, a *mimus*, of the normal attitude toward a hated, disarmed, prostrated enemy."[80] The imagery is important here. The word enemy implies a threat that one fears, but the disarmed enemy is no longer a threat. What occurs during punishment can be merely an attempt to exorcise fear, rather than a reaction to legitimate retributive emotions. As Moore writes, "Yet unlike the victor in a fight who has won and can afford to be merciful to a vanquished foe, those who wish to punish may feel this is their first opportunity to get back, an opportunity they

[79] *Id.*

[80] Nietzsche 1969: 71.

cannot afford to pass up."[81] Expressions of remorse may act as admissions of defeat on the part of criminal offenders that reassure non-offenders of their safety.

Herd Morality

Moore sees more than just fear in Nietzsche's allusion to punishing a captured prisoner. There is also a drive to safety in numbers; to herd morality.

> When Christ talks about throwing stones, it is not because we are equally guilty that we should not throw stones; rather, there is something cowardly in a group of persons throwing stones at one who is now helpless. Such cowardice can be exhibited by the need of such persons for group reinforcement (which is why avengers may refuse to throw the *first* stone—it would set one apart from the group). It is no accident that the retributive urge calls up images of mobs, groups who together finally find the strength to strike back at an only now helpless foe.[82]

However, the urge to be part of the group is not just for protection, but also affirmation of the self and one's morality. "'We good men—*we are the just'*—what they desire they call, not retaliation, but the 'triumph of justice'; what they hate is not their enemy, no! they hate 'injustice,' they hate 'godlessness'; what they believe in and hope for is not the hope of revenge…but the victory of God, of the *just* God, over the godless…"[83]

When one experiences a retributive emotion it may not indicate the actual blameworthiness of its target. Rather it may originate with the urge to feel safety in numbers. It may begin with the desire to be part of the "just" majority over an

---

[81] Moore 1988: 193.

[82] *Id*.

[83] Nietzsche 1969: 48.

"unjust" minority. Again, remorse may act as a reassurance of this deeper ignoble desire. But the retributive emotions can have even darker origins than the desire to identify with the moral.

Sadism

Another deep, important insight of Nietzsche's, which has been adopted by so many prominent philosophers in his aftermath that today it seems self-evident, is that the retributive emotions and punishment may arise from sadistic impulses rather than actual blameworthiness. There are many illuminating, disturbing passages in his writings to illustrate the point, but I will confine myself to the two I find most powerful. Nietzsche refers to Roman times when the punishment for defaulting on a debt was to have the creditor mutilate the debtor's body:

> [T]he creditor could inflict every kind of indignity and torture upon the body of the debtor; for example cut from it as much as seemed commensurate with the size of the debt—and everywhere and from early times one had exact evaluations, legal evaluations, of the individual limbs and parts of the body from this point of view, some of them going into horrible and minute detail…Let us be clear as to the logic of this form of compensation: it is strange enough…a recompense in the form of a kind of *pleasure*—the pleasure of being allowed to vent his power freely upon one who is powerless, the voluptuous pleasure "*de faire le mal pur le Plaisir de le faire*[84]," the enjoyment of violation…The compensation, then, consists in a warrant for and title to cruelty.[85]

[84] Of doing evil for the pleasure of doing it.

[85] Nietzsche 1969: 64.

One may be tempted to contend that our modern, "humane" forms of punishment prevent there from being a sadistic satisfaction in punishment. However, this argument is belied by the fact that victims are permitted to attend an offender's execution and the explosion of documentary accounts on television that chronicle the hard treatment of offenders without any sympathetic narrative.[86]

The sadistic drive reaches its culmination in Christian theology, where the blessed are permitted to look down on the punishments of the damned for eternity. Nietzsche quotes Tertullian delighting in the torture of all his perceived enemies in hell.[87] The real source of our retributive emotions can be a deep-seated, sadistic impulse. Anyone familiar with a sufficient number of judges, district attorneys, and police officers knows that there are some who draw a distasteful amount of pleasure in seeing criminal offenders generally, or in specific instances, humbled. The begging and pleading for the state's mercy might serve their interests.

Self-deception

Underlying all of these foregoing possible sources of retributive emotions is the urge to self-deception. The retributive emotions (vengeance, the urge to punish, to blame, etc.) can be construed as *prima facie* immoral, or at least unflattering. Even more so, when they arise from blatant self-interest in satisfying one of the drives listed above. One way to make them appear as virtues is to deceive ourselves as to

[86] In my opinion, the television shows "Cops" (which has been enormously successful by TV standards) and NBC's "Locked Up" are two good examples of the sadistic gaze that the public enjoys.

[87] Nietzsche 1969: 49-52.

their origin; to locate their source in the offender and his/her desert rather than in ourselves. Those who blame may be "black magicians, who make whiteness, milk, and innocence of every blackness…"[88] In this way, the drive to self-deception reveals itself as but one more source of the retributive emotions that is unrelated to the actual blameworthiness of the offender.

The Danger to Our Own Virtue

As I hope I have made apparent, there is ample textual support in Nietzsche's writings for the contention that the retributive emotions may not be an epistemically reliable source for determining the desert of punishment. In certain places, Nietzsche goes even further and suggests that we ought to avoid the retributive emotions for their effects on ourselves. "Mistrust all in whom the impulse to punish is powerful. They are people of a low sort and stock."[89] And from the Gay Science: "Let us not become darker ourselves on their [offender's] account, like all those who punish others and feel dissatisfied. Let us sooner step aside. Let us look away."[90] "*Looking away* shall be my only negation."[91]

These passages and those which preceded make it very tempting to conclude that Nietzsche would find those inclined to grant mercy on the basis of remorse as evidence that they have somehow overcome the vicious influence of *ressentiment*.

[88] Nietzsche 1969: 47.

[89] Nietzsche 1954a: 212.

[90] Nietzsche 1974: 254.

[91] Nietzsche 1974: 223.

But this is too simplistic a reading of Nietzsche, for he also contends that our drive to spare offenders punishment may derive from unreliable self-interested emotions.

The Unsavory Origins of the Drive to Abstain from Punishing

As Bernard Williams noted, Nietzsche's texts are "booby-trapped" against recovering neatly worded theory from.[92] This is certainly the case with jumping to the conclusion that Nietzsche would abolish all punishment because the origins of our retributivist emotions are murky. For Nietzsche contends that our drive to abstain from punishment may have its origins in the very same muck:

> There is a point in the history of society when it becomes so pathologically soft and tender that among other things it sides even with those who harm it, criminals, and does this quite seriously and honestly. Punishing somehow seems unfair to it, and it is certain that imagining "punishment" and "being supposed to punish" hurts it, arouses fear in it. "It is not enough to render him *undangerous*? Why still punish? Punishing itself is terrible." With this question, herd morality, the morality of timidity, draws its ultimate consequence…The imperative of herd timidity: "We want that some day there should be *nothing anymore to be afraid of!*"[93]

In this passage we see that weakness, fear, herd morality, pity, and self-deception are also at work in blanket opposition to punishment. There is weakness in not managing the will to punish because punishment itself is terrible. There is the desire to abstain from punishment because punishment itself is scary and the desire is to avoid all that can be terrifying. There is the herd timidity, the safety in numbers

[92] Williams 1994: 238.

[93] Nietzsche 1966: 114.

that one is part of the group and no one in the group will be punished. There is also the self-satisfaction of being a part of the group that does not impose punishment, which is characterized as being just. There is pity for offenders, which is not in opposition to sadism, but rather akin to it. Sadism is taking pleasure in having power over and causing pain to the offender. It is in using the offender for self-interest. Pity is taking pleasure in having power over and granting mercy to an offender. As Moore says it is "the elevation of self by pity."[94] Lastly, there is self-deception regarding the reasons for abstaining from punishment.

This is the position of the "good Christian" Venetians in the Merchant of Venice. They spare Shylock his life, but rob him of his entire identity by stripping him of all his possessions. He will no longer be able to act as a usurer, which was not merely an occupation but an affirmation of his religious identity and his pride in opposition to the power structure of Venice with all its discrimination. Shylock prefers death and asks for it. The "mercy" of the Christian Venetians is every bit as destructive and vindictive as the punitive measures Shylock was legally entitled to and sought. The Christians are just self-deluded about their own virtue.

For Nietzsche, any emotion can have its origin in something other than what it consciously claims. True, the retributive emotions may be especially susceptible in this regard, but there is no textual support in Nietzsche to suggest that the drive to abstain from punishment is any more reliable. Moore diagnoses the challenge posed

[94] Moore 1988: 211.

to theorists in the aftermath of Nietzsche: "If one accepts, as Nietzsche did, that both retributive and anti-retributive judgments are often motivated by, or at least expressions of, non-virtuous emotions, where does that leave us? It should leave us asking whether we cannot make our judgments about punishment in such a way that they are not motivated by either set of unworthy emotions."[95] Unfortunately, Moore's response is only to wade back into the same sort of self-deception he just uncovered.

2. Moore's Response to Nietzsche

Moore makes two claims in response to Nietzsche: 1) Even if retributive emotions make us less virtuous, or "darker ourselves," this does not impact the truth of the judgments we reach through those emotions; and 2) our ability to imagine our guilt were we to commit the crime of the offender is a reliable means for testing their blameworthiness, immune from the self-deception identified by Nietzsche. I contend that each of these moves are errors.

In the first place, Moore is making an assumption that there is an independent, objective truth regarding an offender's blameworthiness and that our varying emotions either help us to get to that truth or not. For Nietzsche there is no objective, disembodied truth. "There is only a perspective seeing, *only* a perspective "knowing"; and the *more* affects we allow to speak about one thing, the *more* eyes, different eyes, we can use to observe one thing, the more complete will our 'concept'

[95] Moore 1988: 212.

of this thing, our 'objectivity,' be."[96] However, as Christopher Janaway argues, it is not a mere multiplicity of affects that results in knowledge, but "There must also be those operations upon affects, or upon interpretations, which Nietzsche calls 'having them in one's power,' 'shifting them in and out,' and so on."[97]

Moore's first mistake is now apparent. He construed Nietzsche's critique of the retributive emotions to be that they are unreliable for getting at the objective truth of blameworthiness. Nietzsche's critique is more nuanced than that. The objection to the retributive emotions that emerge only from weakness, fear, herd morality, sadism, and self-deception is that they are relatively few affects which are all akin as constituents of slave morality. Even worse, these affects are often involuntary reactions that make it difficult to shift perspectives and consider the issue of blameworthiness anew. In adjudging blameworthiness in this way, the results will always be skewed towards that of one gripped by *ressentiment.*

Moore compounds his first mistake by claiming that the issue of blameworthiness should be viewed through the lens of our ability to feel hypothetical guilt for the offender's actions. But, of course, guilt is just another affect that, along with the others enumerated above, collectively comprises slave morality. Moore's methodology is the same as the methodology already critiqued by Nietzsche. It is not a better judge of blameworthiness, and perhaps a worse one since it reduces the perspective to that of just one aspect: guilt.

---

[96] Nietzsche 1969: 119.

[97] Janaway 2009: 56-57

If we are to have knowledge of blameworthiness and virtuous attitudes toward expressions of remorse in Nietzsche's schema, then it is our duty to embrace perspectival knowledge as Nietzsche construes it in order to transcend our involuntary retributive reactions grounded in *ressentiment.*

3. Our Duty to Embrace Perspectival Knowledge of Blameworthiness

Nietzsche does not set forth a methodology for arriving at just determinations of blameworthiness through perspectival knowledge. Nevertheless, it is clear that it is not through the rise of retributive emotions based on weakness, fear, herd morality, sadism or pity, guilt, or self-deception alone. Additionally, there are two passages which shed some light on the shift in perspective that may be necessary in order to gain the sort of knowledge about desert which could be the basis of just punishment.

In the first passage, Zarathustra tells a story of a snake bite that occurred while he was asleep under a tree:

> An adder came and bit [Zarathustra] in the neck, so that Zarathustra cried out in pain. When he had taken his arm from his face, he looked at the snake, and it recognized the eyes of Zarathustra, writhed awkwardly, and wanted to get away. "Oh no," said Zarathustra, "as yet you have not accepted my thanks. You waked me in time, my way is still long." "Your way is short," the adder said sadly; "my poison kills." Zarathustra smiled. "When has a dragon ever died of the poison of a snake?" he said. "But take back your poison. You are not rich enough to give it to me." Then the adder fell around his neck a second time and licked his wound.[98]

[98] Nietzsche 1954a: 179-181.

When asked the moral of the story Zarathustra replies: "[I]f you have an enemy, do not requite him evil with good, for that would put him to shame…And rather be angry than put to shame. And if you are cursed I do not like it that you want to bless. Rather join a little in the cursing. And if you have been done a great wrong, then quickly add five little ones: a gruesome sight is a person single-mindedly obsessed by a wrong."[99]

There are many points that could be drawn from this little parable, but chief among them is that Nietzsche does not univocally condemn retributive emotions or punishment. They can be natural and just. What Nietzsche appears to advocate here is that one's perspective (and eventual response/punishment imposed) must be controlled such that blameworthiness can be assessed without becoming "single-mindedly obsessed by a wrong." Blameworthiness must be assessed from a place of strength and without fear of the offender. There should also be a capability to care for the offender, as expressed in the injunction against putting them to shame, such that they are not merely a tool to vent upon through sadism or elevate oneself upon pity. Foremost, there should be an ability to control one's perspective and not be gripped in *ressentiment*.

Nietzsche takes a similar tact in discussing state punishment:

> As the power and self-confidence of a community increase, the penal law always becomes more moderate; every weakening or imperiling of the former brings with it a restoration of the harsher forms of the latter…The justice which began with, "everything is dischargeable, everything

[99] *Id*.

> must be discharged," ends by winking and letting those incapable of discharging their debt go free: it ends, as does every good thing on earth, by *overcoming itself.* This self-overcoming of justice: one knows the beautiful name it has given itself—*mercy;* it goes without saying that mercy remains the privilege of the most powerful man, or better, his—beyond the law.[100]

Once again we have an image of justice from a position of strength, which reaches its climax in the ability to forego punishment. This is not the requirement that punishment is not imposed, or the assumption that it will not be applied in most cases, but the expression that the proper perspective from which to adjudge blameworthiness is one of strength and without fear. If one cannot imagine having mercy on a specific offender, then one is surely in the grip of *ressentiment* to some extent.

The contrast to how remorse currently operates in our system is stark. Even though the modern state is exceedingly strong compared to individuals (and only getting stronger as technological advances make it easier for a minority to monitor and physically control a much, much larger majority), it behaves as though shackled criminal offenders in the courtroom are still a threat to be cowed into submission. The state behaves as though it is afraid to hear what the offender might say, what they might reveal about the state and its supporters. Virtue and justice demand a more confident and humane approach. The question becomes whether there is a way to decouple remorse from the theater of the courtroom and the imposition of violent

[100] Nietzsche 1969: 72.

punishment such that we have some assurances the practice is not just some self-serving vicious satisfaction of the retributive emotions. One where the criminal offender is more of a respected participant than just a vessel for our darker impulses. Emerging trends in restorative justice may meet these demands.

C. A Way Forward or Dead End?

Restorative justice is a broad catchall category for academic work related to reconciliation in political and criminal processes. In general, the work is about political processes that allow victims, perpetrators, and members of society at large to share stories of wrongdoing, consequences flowing from wrongdoing, statements of remorse, updates on progress in the hope of some benefit accruing to participants. There is significant disagreement about the theoretical and empirical support for the many practices that the literature considers. I do not intend to weigh in on these matters, just to consider whether some of the features in this body of work might address the concerns I have raised about the way remorse functions in the current Anglo-American criminal justice system. I remain open to the value of remorse in the political context of the criminal law under certain constraints.

Amnesty

Restorative justice is especially concerned with societies going through transitional phases, often following the fall of an oppressive regime. In these contexts it may not be feasible to punish all of the offending actors who supported the regime while still desiring to document their accounts of their misdeeds. Accordingly,

restorative justice takes up the question of whether amnesties are permissibly granted in order to incentivize participation.

Much of what I have objected to in the preceding chapters is the threat of violence the state imposes on criminal offenders in order to encourage them to express remorse. If there is value in restorative justice hearings and meetings between victims and offenders, then nothing an offender says in these hearings should be used in consideration of punishment. Restorative justice hearings should be completely bifurcated from hearings bearing on punishment. They should be a sort of black box, where offenders are free to give honest accounts without fear of further repercussions. Surely, if there is any merit in these practices it depends on offenders being unafraid to tell the truth.

Voluntariness

Some strains of restorative justice prioritize the interests of victims over those of offenders to an unreasonable extent. For some, if there is a benefit to be derived by the victim in being able to confront the offender that is sufficient justification for the practice. If a criminal offender's participation in restorative justice hearings is not voluntary then these restorative justice hearings are subject to much of the same criticisms I have directed at current practices.

Reparations or Restitution

There is an emphasis on distinguishing reparations/restitution from punishment in restorative justice contexts. I am not sure that this distinction is logically coherent. Reparations are usually financial transfers from offenders to

victims intended in some way to correct the effects of the wrongdoing. Whether the state imposes corporal pain, confinement, or financial harm on criminal offenders they all carry the whiff of punishment. Imposing less of an obligation for financial compensation on offender's who express remorse in restorative justice hearings is also likely subject to much of the foregoing criticism.

D. Conclusion

Emerging ideas about restorative justice, particularly amnesty and voluntariness in the participation of criminal offenders, may offer guidance to new practices that maintain some role for remorse in a modern system of criminal justice. But this movement is inchoate and will have to be guarded against the strong forces that seek to use the remorse of criminal offenders to reinforce existing power dynamics and vent the vicious emotions of society.

CHAPTER V: SINCERE REMORSE VS. PERFORMATIVE REMORSE AND THE APPROPRIATE CRITERIA FOR JUDGMENTS IN THE CRIMINAL LAW

> *"A fundamental premise of our criminal trial system is that the jury is the lie detector. Determining weight and credibility of witness testimony, therefore, has long been held to be the part of every case [that] belongs to the jury, who are presumed to be fitted for it by their natural intelligence and their practical knowledge of men and the ways of men."*
>
> -United States Supreme Court[101]

In this chapter I distinguish my argument from a similar call to eliminate considerations of remorse from criminal proceedings. In Chapters II and III, I put forward arguments based on moral and political rights for why expressions of remorse should not be used as grounds for imposing less punishment on criminal offenders. That argument applies equally to sincere expressions of remorse and insincere opportunistic expressions of remorse, what I call performative remorse. While my argument is novel, this distinction between sincere and performative remorse is not. Some academic commentators have been arguing that human beings are unreliable detectors of sincere remorse (and deception more generally). In their view, sincere remorse is perfectly just grounds for state mercy, but modern proceedings do not have a reliable way of detecting it. Therefore, they argue, modern scientific examinations (*e.g.* fMRI scanning) and expert judgments based thereon should be introduced to either supplement or even replace the judgment of lay judges

[101] United States v. Scheffer, 523 U.S. 303, 313 (1998) *internal quotations and citations omitted.*

and juries. The implication is that considerations of remorse ought to be postponed until we do so.

In section A of this Chapter, I describe the arguments of cognitive scientists along these lines, the evidence they put forward in support of their claims, as well a strain of philosophy in support of this project (which I call the technocratic approach to legal judgments). This approach prioritizes interpretations of remorse, sincerity, and deception that are standardized, generalized, and measurable. This approach requires redefining and reducing those terms, eliminating much of the complexity and richness of the context in which they arose: the churn of everyday life.

In section B, I go on to contrast the technocratic approach to legal judgments with a more democratic approach rooted in natural language. On this view, any competent speaker of a language is capable of the sort of interpretations and judgments inherent in criminal proceedings. Additionally, there may be forms of evidence that are not measurable in ways typical of the hard sciences, but no less reliable because of that.

Finally, in section C, I compare the way these two different approaches differ in handling real legal issues and argue that the natural language approach is much more consistent with the proper goal of the criminal law: rendering a community's conception of justice.

A. <u>The technocratic approach to legal judgments</u>.

Judges and jurors have long been trusted to distinguish sincere remorse from performative remorse, as well as honesty from deception more generally. For

instance, the U.S. Supreme Court has held that a criminal offender cannot be forcibly medicated during sentencing hearings where those medications would interfere with the judge's or jury's ability to "know the heart and mind of the offender," including "his contrition or its absence."[102] The legal paradigm is that judges and jurors are capable of reliably judging sincerity so long as they are permitted to observe witnesses or offenders expressions, mannerisms, tone of voice, etc.

This commitment is so strong that in modern American criminal trials testimony is almost only admissible into evidence when it is given live, in-person, in the presence of the fact finder (always a judge or jury, none of which are likely to have scientific training). Hearsay rules generally prohibit the admission of any out of court testimony offered to prove the truth of the matter asserted in the statement.[103] Even where witnesses are permanently unavailable (say, dead), their previous statements are frequently not allowed into evidence unless they fall under a number of very specific circumstances that serve as an indication of their credibility (for instance a deathbed declaration of the identity of one's killer).[104] There is an overwhelming preference that fact finders be able to observe witnesses as they testify. The practice is expensive and time consuming, and allowing experts to take statements from witnesses and summarize or excerpt them for fact finders would certainly be more efficient. However the commitment to lay judgment through live

---

[102] Riggins v. Nevada, 504 U.S. 127 (1992).

[103] *Fed. R. Evid.* 401-403.

[104] *Fed. R. Evid. 804.*

observation is so strong that many courts postponed criminal trials altogether during the height of the Covid-19 pandemic.[105]

Criticism of the abilities of laypeople to make these sorts of judgments in legal proceedings are as old as the institution itself. The target of much of this criticism is quite often the use of a jury as factfinder. "[T]he jury system…entails frequent miscarriages of justice, permitting thousands of notorious criminals to escape, and disposing of important rights of property not according to law, but in accordance with the prejudice and sympathies of comparatively ignorant men, thus depriving law of what should be one of its most prominent characteristics, namely, certainty."[106] Or as the slightly more modern version goes: "When you go into court you are putting your fate into the hands of twelve people who weren't smart enough to get out of jury duty."[107] These earliest criticisms were often accompanied by calls to abolish the jury in favor of the supposedly more learned, temperate, and accurate judgments of judges.

Modern cognitive scientists continue this tradition of branding juries inaccurate, but they do not stop there. Judges are also no longer thought to be sufficiently reliable either. "Our legal system is one of the most impressive feats of

[105] Whitehurst & Tarm 2022.

[106] Hayne 1884: 353. *See* Stanley 2021 for a much poorer modern retread of the same argument.

[107] This joke is usually attributed to Norm Crosby, a standup comedian who frequented the Carson show in the 70's. But I suspect it has existed in one form or another for as long as the institution it mocks. I personally prefer Robert Frost's quip: "A jury consists of twelve persons chosen to decide who has the better lawyer."

Western civilization. But psychology and neuroscience in recent years have shown many of its tacit assumptions to be out of sync with our best understanding of how our brains and minds work…[C]an a judge or jurors infer a defendant's emotions reliably?…My lab…has…discovered a hitch in this paradigm…People turn out to be quite bad at inferring emotions without context. This includes judges and juries…It's not even clear that judges are reliable arbiters of what *they* are feeling."[108]

Accordingly, attempts are being made to introduce fMRI examinations of witnesses and offenders (combined with expert interpretation of the results) to establish whether deception is present or not.[109]

1. The evidence cognitive scientists present in favor of the technocratic approach.

There is one primary experimental paradigm to support the claim that human beings are unreliable detectors of remorse (or any other emotion): "the mind in the eyes test." In these tests subjects are presented with images solely of the eye region of a face (generally obtained by cutting out images of actors and actresses from magazine photos).[110] Accompanying the image are varying number of emotional ascriptions ("serious; ashamed; alarmed; bewildered") that are sometimes antonyms or sometimes are supposed to have the same emotional valence.[111] These ascriptions

[108] Barrett 2017.

[109] United States v. Semrau, 693 F.3d 510, 6th Cir. (2012).

[110] Baron-Cohen *et al.* 1997.

[111] Baron-Cohen *et al.* 2001.

are generated by the authors and screened by a self-selected panel of judges until they produce something like a consensus. In one study five out of eight judges had to agree "that the target word was the most suitable for each stimulus," and no more than two of the judges could have selected any one of the supposedly erroneous foils.[112] There is no indication that the authors ever reached out to the people whose eye regions are used in the exam to determine what their self-reported mental state was at the time the photograph was taken. Or what sort of emotional state they were trying to convey if they were merely acting or posing in that moment.[113]

In some of the tests supposedly normal people have barely performed above what would be expected from luck. In other versions lay people get as many as seventy percent correct.[114]

At first glance, these experiments appear to tell us more about the abilities of the authors to craft attractive ascriptions (something like marketing) than they do about the population at large or any specific individual's abilities to accurately detect the emotional state of another human being with the context of seeing them speak and observe their mannerisms. There is not an obvious connection between accurately identifying a third party's ascription of the emotion portrayed by a set of eyes and being able to accurately identify the emotion of the person connected to those eyes.

---

[112] Baron-Cohen *et al.* 2001.

[113] An interesting question is whether actors need to try to feel the actual emotions they are trying to portray in order to be convincing; such as in the method acting vs. classical acting distinction.

[114] Barrett 2017.

There may turn out to be some such connection, but there is certainly not sufficient evidence on the basis of these studies alone.

Additionally, these studies have been criticized for bias. They often prime participants by asking questions about the features of the participant before the exam begins (*e.g.* biological sex or gender or being the parent of someone with Asperger's syndrome). Some of these characteristics are associated with either higher or lower than normal ability to empathize within our culture, such that the way the tests are administered might be affecting the performance of some of the individual participants. Finally, the tests have been criticized for plainly not detecting the significant results that the authors claim.[115]

But that has not stopped cognitive scientists from drawing sweeping conclusions about reliability in ascribing emotional states generally: 1) that normal people are quite bad at it;[116] that they are kind of okay at it;[117] that there is a significant difference in ability based on biological sex;[118] and in those with and without Asperger's Syndrome;[119] and in psychopathic individuals;[120] and so on.

In my mind, the cognitive scientists in these cases have been led by their ability to standardize and measure into an illusion of simplistic determinacy. Does

[115] Fine 2011: 17, 18, 108-109, 248-249

[116] Barrett *et al.* 2019.

[117] Baron-Cohen *et al.* 2001.

[118] Baron-Cohen 2004.

[119] Baron-Cohen et al. 1995, 1997, 2001.

[120] Richell *et al.* 2003.

the ability to produce a meager consensus or a lack thereof in the highly unusual context of looking at pictures of eyes and very simplistic ascriptions of emotion really tell us much of about the complex phenomena that is detecting emotions of others? Certainly not with the current state of the research. But even as the research advances, I argue in Section C below that there is reason to pause before adopting the criteria of cognitive scientists and experimental psychologists in the arena of the criminal law. This is because the technocratic approach redefines/reduces what it means to be sincerely remorseful to whatever instances are detectable, measurable, and able to be standardized using their methods. But the proper understanding of these terms in the context of the criminal law is their everyday use and understanding within the society where that law is in effect.

For instance, imagine if we tried to improve our criminal justice system based on the evidence the cognitive scientists have amassed thus far. We might limit our pool of jurors or judges to those who score exceptionally well on the reading the mind in the eyes test so that we have measurable proof that they are reliable detectors of emotion. What we will have done in that case is to prioritize facial expressions in the region of the eyes over whatever other characteristics a given society typically associates with sincere remorse (possible examples: trembling voice, hanging head, saying words like "I'm sorry").

This ability to measure leads the cognitive scientists to believe that progress through technological advancement and expertise is possible when it comes to

detecting sincerity, remorse, and/or deception. They have a strain of philosophy in support of this project.

2. Philosophical support for the technocratic approach.

The question of where the authority resides to interpret and define the meaning of words like remorse, sincerity, and deception is a question of where philosophical authority resides. There is an approach to philosophy of language that would locate that authority in experts and their ability to advance understanding progressively in the way that scientific knowledge is thought to advance. One such thinker is Timothy Williamson.

Williamson's approach to philosophy and language locates the capacity to engage in productive philosophy not in any competent user of language, but in some philosophical training beyond language competency (primarily logic) and acceptance of the existing dominant philosophical paradigm. The upshot is that philosophy is a field governed and constituted by philosophical experts.

Williamson's project is to show that "Although philosophers have more reason than physicists to consider matters of language or thought, philosophy is in no deep sense a linguistic or conceptual inquiry, any more than physics is."[121] To meet his goal, Williamson gives an account of shared language that is not dependent on similarity between the users of that language, but on some causal relation between the users; namely logic.

[121] Williamson 2007: 21.

Williamson builds his account of shared language off of work by Kripke, Putnam, and Burge. Williamson attributes to these philosophers the general observation that different speakers make "asymmetric contributions" to the shared language. For instance, experts wield greater influence (near total influence) over the given meaning of a technical word for non-experts. Williamson takes this general asymmetry as proof that there is no "invariant core of beliefs," "canonical list of 'criteria,'" or "shared stock of platitudes" common to the users of a shared language in order for them to be able to use words with a given meaning.[122] Rather, "Speakers may simply differ from each other in various ways in their ability to distinguish between members and non-members of the relevant kind."[123]

However, this is not to state that language is individualized. There remains a dependence on other speakers for meaning of language. "[I]ndividual speakers defer to the linguistic community as a whole. They use a word as a word of public language, allowing its reference in their mouths to be fixed by its use over the whole community."[124] Which is not to say that the majority of users are the sole determinant of a given word's meaning, but the social meaning will remain primary. "The point of the social determination of meaning is not that meaning can never be

[122] Williamson 2007: 123.

[123] Williamson 2007: 124.

[124] Williamson 2007: 124.

determined individually, but that, when an individual does use a shared language as such, individual meaning is parasitic on social meaning."[125]

In this way, Williamson claims that what makes for a shared language is "the complex interrelations of the constituents, above all, their causal interrelations."[126] For two users to use the same word means that there is necessarily a causal relation between them. Likewise, "Creatures who are causally unrelated to us cannot use our word 'not'; at best they can use a word exactly like our word in its general syntactic, semantic, and phonetic properties. But, on the usual view, their word can in principle be synonymous with ours. Expressions are synonymous when they have exactly the same semantic properties."[127] Williamson goes on to sketch what sameness of semantic properties might be: at least sameness of intension, but meaning is not made up in context or dependent on speakers treating expressions interchangeably.

Williamson is careful to point out that this sketch may not be complete, but he is adamant that the addition of newly identified semantic properties that may be required to establish synonymy will not depend on similarities in actual users. "Whether an expression in one language is synonymous with an expression in another language is not a matter of whether the two speech communities associate similar beliefs with the expressions. Rather, the practices of each community (including their

[125] Williamson 2007: 125.

[126] Williamson 2007: 123.

[127] Williamson 2007: 127.

beliefs) determine the sematic properties of its expressions. Synonymy is the identity of the properties so determined irrespective of similarities in beliefs."[128]

Since meaning, for Williamson, is largely divorced from beliefs/context/culture and can be studied objectively from a distance through semantic properties, Williamson can be comfortable stating that "By ordinary working standards, the word 'synonymous,' is quite clear enough to be useful."[129] For Williamson there is no requirement that meaning fit into a broader schema accepted publicly. It can merely be an asymmetrical contribution by himself and/or linguists to the language.

The asymmetry that exists between experts and other users of shared language is furthered in the practice of philosophy where it becomes apparent that Williamson believes that expert philosophers are in a superior epistemic position. Williamson illustrates his point through a discussion of skeptics.

Williamson says that judgment skepticism is doubting or questioning "our practices of applying concepts in judgment." However, total judgment skepticism (that is to say skepticism about all judgments) is rare because it results in intellectual paralysis. Instead, most judgment skeptics tend to be skeptical "about some contextually relevant judgments."[130] Often, judgment skeptics are in apparent support

[128] Williamson 2007: 128.

[129] Williamson 2007: 50.

[130] Williamson 2007: 220-224.

of science or advances in metaphysics as superior to ordinary, common, or folk conceptual practices.

Williamson contrasts judgment skepticism with what he calls traditional skepticism. Traditional skeptics tend to have broader targets than judgment skeptics. Traditional skeptics might question perception, memory, testimony, or the existence of the external world. Additionally, a traditional skeptic is more likely to argue that we cannot know whether we are in a generally assumed scenario (e.g. a brain in a human body) or a skeptical scenario (e.g. a brain in a vat). But a judgment skeptic is more likely to argue that we are actually in their skeptical scenario (e.g. we use a concept that may be in some sense practically useful, but the practice is ultimately without support in natural science).

However, even when judgment skeptics take this affirmative position, Williamson claims that the underlying structure of their argument mirrors that of the traditional skeptic. "In each case, the skeptic concedes an evidential base, in order to accuse us of going illegitimately beyond it." The traditional skeptic may concede that we have the sensation of perceiving mountains, but argue that these sensations cannot be proven to be actual perceptions of mountains such that we cannot know and are not justified in believing that there are mountains. Likewise, the judgment skeptic may concede that we have and apply the concept "mountain," but argue that there is not truth in the application of the concept because what we call a mountain is really only a complex microphysical event. The general form of the judgment skeptic's argument means that it often calls into question much more than the judgment skeptic

wishes, including the natural sciences. "In particular, judgment skeptics who judge that our empirical evidence tells against the reliability of some folk theory are vulnerable to judgment skepticism about the elements of folk epistemology on which they are relying."[131]

For Williamson, evaluating the claims of judgment skepticism turns on evaluation of evidence. Williamson claims that philosophy is a discipline dependent on evidence to no less of an extent than math or the natural sciences.[132] Williamson describes this as a pragmatic view of evidence whereby "what permits a fact to serve as evidence in a given context is that it happens to be uncontroversial in that context."[133] Williamson claims that this model of evidence works well enough when engaged in dialog with those who show "moderation and restraint" in declaring evidence controversial. The problem with judgment skeptics is that they challenge too much of the possible evidence. "[F]or one regards their restricted evidence base as too willfully impoverished to constitute a reasonable starting-point for inquiry." Skeptics who go the furthest violate Carnap's "requirement of total evidence." In philosophy, it is standard practice to tolerate skepticism and even to engage skeptics in conversation, but one is not obligated to take their points as epistemically relevant.

[131] Williamson 2007: 224.

[132] Williamson 2207: 208, 241, 246.

[133] Williamson 2007: 238.

"There is no bad faith in continuing to claim (and have) knowledge of the contested truths…Sometimes, in self-defense, one must abandon skeptics to their fate."[134]

The epistemic asymmetry I referred to above occurs in determining when one's interlocutor is exercising appropriate "moderation and restraint" to allow for an epistemically fruitful dialectic. It seems that Williamson believes that individuals or select groups have an obvious/clear/decisive way of determining what relevant evidence is. For example, discussing scientists Williamson states: "When scientists state their evidence in their publications, they state mainly non-psychological facts (unless they are psychologists); are they not best placed to know what their evidence is?"[135]

Setting aside the issue of whether Williamson has a coherent account of the obvious/clear/decisive way experts have of determining what relevant evidence is and when their interlocutors are impermissibly narrowing the field of evidence, what emerges is that Williamson's account of both language and philosophy are essentially top-down endeavors. Experts make asymmetrical contributions to both language and philosophy. Experts determine what the relevant evidence is for the meaning of a word and what sort of arguments are reasonable in philosophy, and thus they also control where explanatory power resides. The capacity to practice philosophy, that is to have one's arguments taken seriously, is dependent upon one's acceptance of this model and its methods.

---

[134] Williamson 2007: 239-240.

[135] Williamson 2007: 212.

Williamson's goal with this account is quite clear. He wants philosophy to progress as the "hard sciences" are believed to. "[I]f the philosophical community has the will, it can gradually bring up a much higher proportion of practice to the standard of current best practice, and beyond. Such progress in methodology cannot be relied on to happen automatically; not all of us love the highest at first sight. Although the envisaged incremental progress lacks the drama after which some philosophers still hanker, that hankering is itself a symptom of intellectual immaturity that helps hold philosophy back."[136]

3. Williamson's account of language and philosophy support the cognitive scientists' project.

Williamson, and philosophers like him, are in full support of projects like the technocratic approach to legal judgments. In their view, while words like remorse, sincerity, and deception arose within a specific social context their correct meaning is not dependent on that context. Philosophical judgments about meaning are not linguistic. Rather, they turn on empirical determinations. Their meaning is asymmetrically determined by experts who are properly situated to determine the various causal connections. The privileged epistemic position of experts gives them authority over their meaning. Thus it is permissible that the meaning of concepts like remorse, sincerity, and deception are altered from their everyday usage. As their research project advances, cognitive scientists will continue to alter the meaning of

[136] Williamson 2007: 8.

these terms such that they conform to the evidence. This is the sort of incremental progress of science and technology that we should expect and welcome. And along the way, objections over meaning will be politely ignored if the experts do not determine that they contribute to their epistemic progress. In this way the experts are allowed to self-select and justifiably exclude laypeople. This process is what I mean by the technocratic approach to legal judgments.

B. <u>The natural language approach to legal judgments</u>.

There is an alternative approach that preserves the role of laypeople in legal judgments as well as the ordinary meaning of concepts like remorse, sincerity, and deception. I will describe this account primarily through the work of Stanley Cavell, before introducing Wittgenstein's theory of imponderable evidence as an alternative to the technocratic requirement that evidence be measurable and determinate in every arena.

1. <u>Natural language philosophy and attunement.</u>

Cavell's conception of philosophy locates the capacity to engage in productive philosophy (and more generally to have one's normative judgments taken seriously absent philosophical training) in every competent user of language within the community. Cavell approaches philosophy from the tradition of the so-called ordinary language philosophers. Ordinary language philosophers argue that knowledge is best (perhaps only) understood as derived through analysis of ordinary language. As the ordinary language philosopher J. L. Austin states: "[O]ur common stock of words embodies all the distinctions men have found worth drawing, and the

connexions they have found worth marking, in the lifetimes of many generations: these surely are likely to be more numerous, more sound, since they have stood up to the long test of survival of the fittest, and more subtle, at least in all ordinary and reasonably practical matters, than any that you or I are likely to think up in our armchairs of an afternoon—the most favoured alternative method."[137]

In Cavell's version of ordinary language philosophy, this claim to epistemic authority in ordinary language is conceived of through the concept of attunement. In the most general terms, attunement appears to be agreement over the use of language in given contexts (referencing Kant, Cavell sometimes refers to it in its entirety as "the schematism"). The two most important features of Cavell's idea of attunement, at least for the purposes of this paper, are its pervasiveness and its connection to reason.

For Cavell, attunement implies that our agreement over language is not limited to certain, discrete occasions. Rather, "the idea of agreement here is…being in agreement throughout, being in harmony, like pitches or tones, or clocks, or weighing scales, or columns or figures." In their shared language, human beings are "mutually voiced with respect to it, mutually *attuned* top to bottom."[138] The depth of the pervasiveness of attunement is exemplified well in an example of the concept of a chair, which Cavell borrows from Wittgenstein. Cavell states that it is part of the grammar of the word chair that one uses it in a certain way: by sitting on it. But even

[137] Austin 1956-57: 8

[138] Cavell 1979: 73.

more specifically, one sits on it in a particular way, different than one can and does sit on different objects. And yet, even where one can sit on something in the same manner as they do a chair, one often recognizes that that thing is not a chair, but merely being used as a chair. Attunement is being alive to these many distinctions and nuances across the entire language.

Cavell also identifies the search for shared agreement in our language with a claim to community. Community, in the sense Cavell is using the term, carries the connotation of a shared sensibility, such that understanding and comprehension is possible between members. "I have nothing more to go on than my conviction, my sense that I make sense…The wish and search for community are the wish and search for reason."[139] The idea of community is essential to attunement because it is the basis upon which mutual understanding is built. Specifically, Cavell explores how attunement is dependent on specific contexts that occur within our form of life: "language could not function as it does without a mutual and common agreement about *what* is being named or pointed to. And this depends on our sharing a sense of what is remarkable, or on our attention being drawn in similar directions by similar occurrences; depends upon these in as fundamental a way as it depends on our having similar capacities of sense and action. And it depends upon a sense of what claim will have a point in certain contexts, and a knowledge of what the point is."[140] Now words can be projected into new contexts in a way that is natural or ordinary. But the

[139] Cavell 1979: 20.

[140] Cavell 1979: 211.

danger is when they are projected into contexts that are neither natural or ordinary. "What is left out of an expression if it is used 'outside its ordinary language game' is not necessarily what the *words* mean (they may mean what they always did, what a good dictionary says they mean), but what we mean in using them when and where we do. The point of saying them is lost…What we lose is a full realization of what we are saying; we no longer know what *we* mean."[141]

The upshot of this is that all of our knowledge is tied directly to our form of life. That is, what we have knowledge of is what is mutually remarkable to ourselves and others in our community. Our ability to express that knowledge through judgments is also dependent on our being understood by other members of the community. Thus it appears that for Cavell the existence of attunement (our ability to understand each other) is the basis of our ability to reason at all. Locating the ability to reason at this basic level of what is common throughout the community is what ensures that each competent user of language within the community is capable of philosophy under Cavell's account.

That Cavell grants equal epistemic status to every competent user of language within a community is most evident in his treatment of disagreement about what is appropriate to say in specific contexts, cases, or examples. The idea of attunement is not that disagreement is impossible, but that it is rare and usually derives from an important or interesting source which may be tied to the individual (such as cultural

[141] Cavell 1979: 207

differences, philosophizing, or withdrawing from attunement through skepticism) or to the particular case at hand. Under Cavell's approach to philosophy, disagreements are not resolved by showing that one side was in a superior epistemic position and the other inferior, but rather by analyzing the disagreement in order to reveal some philosophical insight. Cavell's approach to skeptics is a useful illustration.

In Part II of *The Claim of Reason*, Cavell starts with the observation that skeptical inquiries tend to begin with the consideration of generic objects under what are generally considered the best case circumstances for our knowledge of objects. Cavell notes that in these situations, the question of whether we know the object does not arise in a fully natural way. Rather there is something odd about questioning a claim to knowledge under such circumstances, and this oddness suggests a new direction of inquiry: "whence comes this sense of something amiss about the simplest claim to knowledge under optimal conditions, where there is no practical problem moving us?"[142] Cavell's answer is that this sensation is not produced by the philosopher's investigation. Instead, "the philosopher *begins* his investigation with the *sense* that…something is, or may be amiss with knowledge as a whole."[143] Thus, the "phenomenological form of the investigation is, after the fact, that of having confirmed our worst fear for knowledge."[144] The remainder of Part II is an analysis

[142] Cavell 1979: 139-140.

[143] Cavell 1979: 140.

[144] Cavell 1979: 145.

of what makes this skeptical fear possible: 1) an ability to not know the meaning of what we say; and 2) a dissatisfaction with the human conditions of knowledge.

It is an awareness of the contingency of our knowledge which leads the skeptical philosopher to use words outside of their ordinary language games. "The philosopher feels that he must say and think beyond these conditions; he wants to speak without the commitments speech exacts…In philosophizing we come to be dissatisfied with answers which depend upon *our* meaning something by an expression, as though what *we* meant by it were more or less arbitrary."[145] This dissatisfaction is what Cavell at times calls philosophy's "rejection of the human." This dissatisfaction gives rise to an urge for certainty, and a form of knowledge that is not contingent on our human perspective. This is the sense that initiates skeptical investigations.

Thus, we might state that the skeptic is not wrong *per se*, and the non-skeptic is not wrong in any usual way. Nor is a disagreement between a skeptic and non-skeptic going to necessarily dissolve after such analysis. It simply cannot be said that either the skeptic or the non-skeptic is an inferior epistemic position. However, by analyzing the disagreement much can be learned about our type of knowledge and our level of comfort in relation thereto. It is for this reason that Cavell says that philosophy is appropriate, "When…you need a clear view of what you already

[145] Cavell 1979: 215.

know."[146] The nature of a philosophical insight is to reveal something that we already know but that we do not realize for one reason or another.

One objection to the democratic interpretation of philosophy which I claim Cavell endorses, is that philosophy remains an esoteric subject practiced by a relatively small number of human beings. The objection might be levied like this: Cavell's own work is so dense, the structure of its sentences so intricate, the vocabulary so scholarly, the references so broad, etc. that surely it is inaccessible to all but a small group of experts.

I think that one way the objection can be dealt with is by keeping in mind the distinction between capacity and the urge to utilize it. "If philosophy is esoteric, that is not because a few men guard its knowledge, but because most men guard themselves against it."[147] Cavell implies that the capacity for philosophy is broad, but the urge to utilize that capacity is relatively limited. "The effort is irrelevant and worthless until it becomes necessary to you to know such things. There is the audience of philosophy; but there also, while it lasts, is its performance."[148] Thus, the capacity is extensive, the urge is narrow, and the performance will be highly individualized and broadly or narrowly accessible depending upon the philosopher.

But I want to urge that there is a better way to address this objection, which is to reject its strict limitation on what it is to do philosophy (writing books, etc). I will

---

[146] Cavell 1958: 21.

[147] Cavell 1958: xxvii.

[148] Cavell 1958: xxix.

argue in Section C that philosophy is practiced widely and often. In particular, in determining what to say in particular political and legal cases. In this way, judges and juries can be understood to be practicing philosophy when they render certain kinds of judgments in legal cases. In fact, these are the sort of judgments that laypeople are especially well-suited for.

2. Imponderable evidence: a reliable source of knowledge that is, in a sense, immeasurable.

Cavell's work above provides an alternative theoretical basis to counter the technocratic approach's contention that experts will necessarily be better positioned than laypeople to distinguish between sincere remorse and performative remorse. In this section I will introduce Wittgenstein's conception of imponderable evidence as a source of knowledge that is an alternative to the technocrat's demand that evidence be measurable and standardized.

Given that much of Wittgenstein's later work cautions against arbitrarily sharp definitions of concepts[149], it would be strange to attempt a straightforward definition of imponderable evidence. Perhaps even more so since, as Ray Monk notes, "The notion of 'imponderable evidence' is somewhat slippery, and there are signs that Wittgenstein himself was, at times at least, somewhat sceptical of it."[150] This part of the paper will instead gesture at some of features of imponderable evidence that

[149] Cf. Wittgenstein 2009: *Philosophical Investigations* Nos. 35, 51, 52, 69-75, 77, 79, 80, 84, and 87.

[150] Monk 2005: 102.

Wittgenstein seems to think set it apart. In particular, the way that imponderable evidence persuades, how one recognizes it as evidence, and the sorts of judgments imponderable evidence appears particularly useful to are all features relevant to modern criminal trial practice.

As an entry point into understanding what Wittgenstein means by "imponderable," Monk considers the following passages:[151]

> 922. I tell someone that I have reasons for this claim or proofs for it, but that they are 'imponderable.'
>
> Well, *for instance*, I have seen the look which one person has given another. I say "If you had seen it you would have said the same thing." [But here there is still some unclarity.] Some other time perhaps, I might get him to see this look, and then he will be convinced. That would be *one* possibility.
>
> 925. An important fact here is that we learn certain things only through long experience and not from a course in school. How, for instance, does one develop the eye of a connoisseur? Someone says, for example: "This picture was not painted by such-and-such a master"—the statement he makes is thus not an aesthetic judgment, but one that can be proved by documentation. He may not be able to give good reasons for his verdict.—How did he learn it? Could someone have taught him? Quite.—Not in the *same* way as one learns to calculate. A great deal of *experience* was necessary. That is, the learner probably had to look at and compare a large number of pictures by various masters again and again. In doing this he could have been given *hints*. Well, that was the process of *learning*. But then he looked at a picture and made a judgment about it. In most cases he was able to list reasons for his judgment, but generally it wasn't *they* that were convincing.
>
> 927. A connoisseur couldn't make himself understood to a jury, for instance. That is, they would understand his statement, but not his reasons. He can give intimations to another connoisseur, and the latter will understand them.

[151] Wittgenstein 1982: 115.

From this passages, Monk draws the conclusion that imponderable evidence has the characteristic that "it can be *seen* as evidence for a particular judgment, but usually it cannot be described other than as evidence for that judgment (e.g., 'How do you *know* your father dislikes your boyfriend?' 'I could tell by the way he looked at him' 'And how did he look at him?' 'Well,...as if he didn't like him')."[152]

Monk's conclusion is unsatisfactory. Or at least the way he has phrased it is. It just seems slightly at odds with the passages he cites. For instance, the art expert in 925 is perfectly "able to list reasons for his judgment." And the connoisseur's statement in 927 can be understood by the jury, even if they do not understand his reasons. More importantly, another connoisseur could understand them. Also, Monk's interpretation runs contrary to how we expect people to behave when describing imponderable evidence. Taking Monk's example, when asked to describe the father's disdainful look at the boyfriend, would the young person really be at such a loss for words? Do we not try to describe imponderable evidence all the time? "Well his brow was knit and he never smiled. Also he was squinting his eyes." Wittgenstein is certainly saying there is something problematic with descriptions of imponderable evidence, but not that it cannot be described.

In my opinion, the better interpretation of these passages is that Wittgenstein's claim is not that the evidence cannot be described in words, but that its persuasive

[152] Monk 2005: 104.

power does not lie in the description of measurable observations.[153] "Imponderable evidence includes subtleties of glance, of gesture, of tone."[154] In some sense, we can imagine a perfectly accurate, exhaustive description of tone that included decibel fluctuations, pitch, etc. And yet reading that description would not persuade one as to whether the tone was sincere or not. The same goes for a painting where the paint in a certain place might be x centimeters thick, of a hue that produces such and such a reading on a spectrophotometer, and with brush stroke grooves exactly this deep and that wide. And yet that description would not persuade one that the piece was a Rembrandt. Why might this be so?

One way to put it might be that imponderable evidence does not persuade us consciously. The persuasive power is in experiencing the evidence, not merely thinking about it. It can be put into words (described), but the words will only be persuasive for someone who has the ability to integrate these details with their experience of similar contexts. For instance, describing a father's disdainful look ("squinted eyes, knit brows, straight mouth") might not persuade a friend that has never met your father that he disliked your boyfriend. The description might be more consonant with the friend's experience of looks that mean concentration or resolve. But the description might persuade your sibling, who could rely on your description to conjure up their own experience of the fathers' face. This seems to be the way one connoisseur can understand another. Familiarity and the ability to attend to the

[153] Perhaps this is what Monk means as well, but it is not obvious.

[154] Wittgenstein 2009: 240.

appropriate details dependent on the context are crucial to reliably drawing these inferences.

Evidence is ponderable when its persuasive power must appeal to us consciously. For instance, when a lawyer tells the jury that a witness has been promised a reduction in their sentence for testifying in support of the prosecution there is nothing imponderable going on. The jury is being asked to consciously infer that self-interest rather than commitment to the truth is motivating the witness's testimony. The evidence of self-interest is persuasive only if the jury consciously follows that inference. The inference is something like a rule. Logic or reason seem to be more of the guide to how we ought to draw these inferences.

Thus, imponderable evidence persuades without reflection, deliberation, or direct appeal to logic. In this sense, imponderable evidence appears to be consonant with our legal systems insistence that witnesses prevent their evidence, and offenders express their statements prior to sentencing, live and in court where they can be observed. Reading a transcription of just the words would not be sufficient.

The sense in which Wittgenstein believes that imponderable evidence acts as evidence also appears to be consistent with modern legal practice.

> Evidence means testimony, writings, material objects, or other things presented to the senses that are offered to prove the existence or nonexistence of a fact.[155]

[155] *Cal. Evid. Code* § 140.

> Evidence is relevant if: (a) it has any tendency to make a fact more or less probable than it would be without the evidence; and (b) the fact is of consequence in determining the action.[156]

As the two laws quoted above demonstrate, the concept of evidence in American legal proceedings is broad, inclusive, and relatively straight forward. Evidence is basically anything whatsoever that makes some fact more or less likely. This seems to be consistent with Wittgenstein's conception of imponderable evidence. Imponderable evidence is only evidence to the extent that it makes some ponderable fact more or less likely. As Wittgenstein writes:

> 359. The question is: what does imponderable evidence *accomplish?*
>
> Suppose there were imponderable evidence for the chemical (internal) structure of a substance; still, it would have to prove itself to be evidence by certain consequences which *are* ponderable.
>
> (Imponderable evidence might convince someone that a picture was a genuine…But this *may* be proved right by documentation as well.)[157]

Worth noting is that imponderable and ponderable evidence are not mutually exclusive. At least some facts may be proven by either form of evidence. Additionally, imponderable evidence is not infallible, nor even always preferable. Art experts can error in their attributions. But, of course, documentation can also be vague, forged, lost, mistaken, etc. However, judging by the examples Wittgenstein

[156] *Fed. Rule of Evid.* § 401.

[157] Wittgenstein 2009: 240.

marshals, it appears that imponderable evidence is particularly useful in certain areas; i.e. those that are tied most closely to human experience, to our form of life.

There are three general examples that Wittgenstein turns to in his writings on imponderable evidence: 1) art[158]; 2) the sincerity of emotional displays[159]; and 3) the meaning of words.[160] The choice of these examples alone demonstrates that Wittgenstein believes the case for imponderable evidence is strongest in those areas where humans have the most experience. Determining the emotions of another and the meaning of words are things every typical human being has done billions of times by the time they reach maturity. Even though Wittgenstein usually confines his remarks on imponderable evidence in connection to art to experts/connoisseurs that have cultivated excess experience with art, aesthetic judgments generally are also something that every typical, mature human has made innumerable times.

Following Ray Monk and Michel Ter Hark, it appears that experience is crucial to the use of imponderable evidence in two distinct ways. The first is identified by Monk as the ability to perceive imponderable evidence at all. Monk writes:

> An inner process stands in need of outward criteria' runs one of the most often quoted aphorisms of *Philosophical Investigations*, an aphorism that many have cited in support of the notion that Wittgenstein was some sort of behaviourist, an interpretation that needs to be resisted. One way of resisting it is to realize

[158] Wittgenstein 1980: 47 (*RPP* 243); Wittgenstein 1982: 115 (*LW I* 925); Wittgenstein 2009 (*PPF* 359).

[159] Wittgenstein 1982: 110-125 (*LW I* 917, 923, 936, 937); Wittgenstein 1992: 87 (*LW II* p. 95); Wittgenstein 2009: 239-240 (*PPF* 352-358, 360-364).

[160] Wittgenstein 1980: 49 (*RPP I* 243).

> what an emphasis Wittgenstein placed on the need for sensitive perception of those 'outward criteria' in all their imponderability.[161]

If I follow Monk correctly, then he reads Wittgenstein as claiming that our ability to detect imponderable evidence and use it reliably is tied to our immersion in complicated social contexts; the weave of human life. Our ability to judge human sincerity flows naturally from the fact that we have interacted with, learned from, and depended on other humans from birth. Accordingly, we have learned that a change in hair color is usually (but not always!) irrelevant to determining sincerity, but a change in the color of another's face might be relevant to whether they are truly angry or are pretending. Monk's reading is consonant with the passages where Wittgenstein discusses how one learns how to use imponderable evidence.[162]

---

[161] Monk 2005: 102.

[162] For instance, *LW I* 925; where Wittgenstein stresses that an expert learns to use imponderable evidence not from school, in the way one learns to calculate. But instead by exposure to "a large number of pictures." (Wittgenstein 1982). In conversation, Janette Dinishak has pointed out to me the similarities between this passage and Aristotle on moral education. "Virtues…we acquire, just as we acquire crafts, by having first activated them. For we learn a craft by producing the same product that we must produce when we have learned it; we become builders, for instance, by building, and we become harpists by playing he harp." (Aristotle 2009: 18-19). These passages also call to mind Ryle's distinction between knowing how and knowing that. (Ryle 1945-46). One might think then that what the cognitive scientists are doing is merely detecting and describing certain features of detecting sincere remorse vs. performative remorse, without in any way freezing in place a correct technique. Just like someone need not know about inertia to ride a bike, they need not consciously think about cognitive science to reliably read emotion. For instance, one might be told that to detect emotion humans look at the expression of the eye region. And might be told about some of the characteristics there (e.g. squinting, furrowing, widening, dilating, etc.). But detecting those changes and interpreting them correctly requires time, practice, and examples. And just like proper technique for riding a bike changes with the terrain and weather, the technique

Michel Ter Hark argues that it is not merely immersion in a form of life that creates the ability to detect and use imponderable evidence. He also reads Wittgenstein as claiming that the targets of judgments where imponderable evidence is useful are defined by that context as well. As Ter Hark writes:

> By introducing his notion of patterns in life Wittgenstein attempts develops [*sic*] an alternative to a referentialistic account of the meaning of psychological concepts. In particular, when Wittgenstein says that, say, 'grief' describes a pattern in our life, he explicitly differentiates ascriptions of grief from, rather than likening them to, sentences describing either bodily behaviour or physiological states and processes. The dependency of psychological concepts upon patterns of life implies that their use is governed by a loose and shifting cluster of descriptions. These clusters may lack a definite or determinate sense.[163]

Given this reading of the targets of certain judgments, ponderable evidence with its use of rule-like conscious inferences is at a serious disadvantage. If the facts themselves are but a pattern in an ever changing weave of life, then abstract rules will quickly find themselves either out of date or not appropriately attendant to the specific context that the judgment concerns. It's useful to think here of slang, which varies by region, generation, class, and so on, but is rarely fixed for long enough that meaning can be pinned down through empirical observation before it has evolved into something new. "Optics" means something very different in a physics lab than it does in a corporation's public relations department.

---

for detecting emotions is also context dependent.

[163] Ter Hark 2000: 211.

3. The natural language approach preserves a role for laypeople and a modes of knowledge that are not easily measured or standardized.

The natural language approach is a near opposite of the technocratic approach. Meaning is generated from the bottom up, as any individual can make contributions to the language to the extent that they can get other speakers to follow along with them. No one speaker is in a better position to judge meaning, because every speaker derives their authority from their participation in the human form of life. The meaning of words shift quickly as do the patterns in that form of life that they refer to. Empirical observation without participation in the process is useless. And participation in the process does not require consciousness of any abstract rules. The evidence that these judgments rely on is not measurable, generalizable, or determinate in the way that the technocratic approach demands. But the evidence and the judgments it leads to are not unreliable because in a significant sense they cannot be separated from each other.

C. The natural language approach to legal judgments is more consistent with the legitimate aim of the criminal law.

Setting aside the question of whether it is Williamson or Cavell's account of language and philosophical authority that is "correct," and I am not sure it even makes sense to ask that question at all, it is fair to wonder which approach is more useful in the context of the criminal law. That is to say, which approach helps the criminal law to achieve its legitimate end: rendering the community's understanding of justice.

1. The technocratic approach violates the principle of legality.

There are several reasons to prefer the natural language approach. First, the approach is more consistent with general claims about the legitimacy of criminal law. These are sometimes referred to as the principle of legality and include precepts such as "judges should not create new crimes, that the criminal law may operate only prospectively, that crimes must be defined with sufficient precision to serve as a guide to lawful conduct,"[164] including factors that will be considered for enhancement or forbearance of punishment. Proponents of the technocratic approach to the criminal law could persuade legislators to pass laws mandating that terms in the criminal codes are to be henceforth interpreted by experts in accordance with the latest evidence from cognitive science. That would likely overcome the hurdle of posed by the first two precepts above. But I do not believe that this would address the spirit of the third precept: whereby the law is supposed to be clear guide for lay citizens. Not knowing that something is illegal is generally not a valid defense to criminal law breakers. However, courts have held that where the law is sufficiently inaccessible to the public there is concern that it denies due process. "Were it otherwise, the evil would be as great as it is when the law is written in print too fine to read or in a language foreign to the community."[165] If lay citizens were forced to follow developments in cognitive science in order to interpret the words of the criminal law there would be little to no chance of the law serving as an action guide.

[164] Kadish & Schulhofer 2001: 293-294.

[165] Lambert v. California, 355 U.S. 225 (1957).

An example that illustrates the how technocratic understanding of credibility differs from lay understanding is the lie detector test.

In order to illustrate what is at stake between the two approaches it is helpful to look at an example that has received attention from the courts: lie detector tests. At the 1933 Chicago World's Fair, Professor Leonard Keeler of Northwestern University's School of Law introduced the first so-called lie detector test. In 1935, the Polygraph examination, as the test was known, was admitted into evidence in a criminal for the first time. However, as the epigraph to this paper shows the Polygraph examination is now nearly completely excluded in American criminal trials, largely due to its unreliability. But that has not stopped those who believe that deception/credibility can be determined via rule-based, scientific processes. Rising to take the place of Professor Keeler and his Polygraph is a new generation of experts/scientists touting EEG (also known as P300) and/or fMRI as lie detector tests that should be admissible in criminal trials. The arguments for and against these technologies will likely turn on whether or not these so-called lie detectors are accurate or not. But, perhaps, the question that should be at the center of the discussion is what does accuracy mean in the context of credibility and whether that can be determined via rule-based, scientific processes at all.

Somewhat surprisingly, Wittgenstein addresses this topic in his later writings:

> And now the question remains whether we would give up our language-game which rests on 'imponderable evidence' and frequently leads to uncertainty, if it were possible to exchange it for a more exact one which by and large would have similar consequences. For instance, we could work with a mechanical

> ‘lie detector’ and redefine a lie as that which causes a deflection on the lie detector.
>
> So the question is: Would we change our way of living if this or that were provided for us?—And how could I answer that?[166]

In this passage, Wittgenstein appears to suggest that uncertainty over the credibility of others is an inherent pattern in our weave of life. Certainly we could eliminate that particular pattern if we all agree to define credibility by some scientific, rule-based mechanism, but to do so would be to change something fundamental about how we live. What that difference might mean is impossible to say. What is clear is that the mechanical lie detector has not made more accurate judgments about credibility so much as changed what credibility means altogether.

Courts are starting to attend to this distinction. In a ruling by the 6th Cir. Court of Appeals upholding the exclusion of expert testimony regarding credibility based on fMRI readings, Judge Stranch wrote, “[T]here are concerns with not only whether fMRI lie detection of ‘real lies’ *has* been tested but whether it *can be tested.*”[167] The fMRI research and testimony at issue in that case, like all so-called lie detector technologies, purported to show only whether an individual was engaged in deception or not (and, of course, whether it had been proven to be able to reach this relatively modest result with sufficient reliability was in dispute). However, determinations of credibility turn on much more than the binary question of whether

[166] Wittgenstein 1992: 95.

[167] United States v. Semrau, 693 F.3d 510, 522 6th Cir. (2012). *Emphases in original*

deception is present or not. The degree and type of deception matter. There are outright lies, but also half-truths, exaggerations, and embellishments. There are purposeful, but also negligent, omissions of details in testimony. There is deception intended to fool observers, but also self-deception so deep that it warps into a form of sincerity. There is self-deception, bias, and self-interest. There are also deceptions that turn out, unintentionally, to be a sort of credible testimony (as when my toddler daughter claims that my immobile infant son broke the lamp across the room). And of course even these patterns are not static within the weave of life, but depend on attending to particular details associated with context.

The point being that the ordinary human concept of credible testimony is much richer than determining whether deception is present or not. In some sense, answering that binary question accurately is largely useless without corresponding judgments related to all of the nuances related to deception in the pattern of human life, only some of which I articulated above. Thus, the credibility of human testimony appears to be exactly the sort of judgment that imponderable evidence is suited to. Human beings, on the other hand, attend to all of these nuances in their everyday interactions.[168] These past experiences attune humans to pick up on

[168] Of course, technologies will continue to advance and become more complex such that they better approximate normal human judgment in some ways. However, as I argue in the next section, when the facts to be interpreted/judged are community established norms then it is members of the community that are best positioned to determine the facts of the matter. In this way, I am very skeptical that technology could advance to the position where it can attend to the intricacies of the weave of human life sufficiently.

imponderable evidence of reliability/deception. Additionally, the target of this judgment (human credibility) is itself a pattern in the human form of life, understood in its full-blooded sense only from within that perspective. Accordingly, there is no reason to suppose that experts aided by technology are at any advantage over a group of typical humans in determining the credibility of witnesses.

2. The natural language approach better helps the criminal law guard against tyranny.

I have stated on occasion that the legitimate purpose of the criminal law is to render a community's sense of justice. Correspondingly, one of the illegitimate uses to which the criminal law is put is tyranny, which is the use of the state's violence to enforce the existing power structure rather than render justice for the community. One justification for the inclusion of laypeople in the criminal justice process has always been that they will somehow act as a bulwark against tyranny.

> The guarantees of jury trial in the Federal and State Constitutions reflect a profound judgment about the way in which law should be enforced and justice administered. A right to jury trial is granted to criminal defendants in order to prevent oppression by the Government. Those who wrote our constitutions knew from history and experience that it was necessary to protect against unfounded criminal charges brought to eliminate enemies and against judges too responsive to the voice of higher authority. The framers of the constitutions strove to create an independent judiciary but insisted upon further protection against arbitrary action. Providing an accused with the right to be tried by a jury of his peers gave him an inestimable safeguard against the corrupt or overzealous prosecutor and against the compliant, biased or eccentric judge.[169]

[169] Duncan v. Louisiana, 391 U.S. 145 (1968)

The protection of the jury is most often thought of in simplified material terms: lay jurors are members of the public and not employees of the state so they do not have the same literal interest in maintaining the existing power of the state as, say, a judge does. But I want to suggest that there is a deeper philosophical way in which the jury guards against tyranny. In liberal democracies sovereignty is located in citizens and then conferred periodically to elected officials. The inclusion of lay judgment in criminal trials is a way of ensuring that the community's interpretation of justice is prioritized in the instances where it is most crucial.

I stated above that I would argue that philosophy is practiced widely and often and I think that an example is sufficient to prove the point. The Judicial Council of California produces model jury instructions that are given to jurors in almost every jury trial in the state. One unremarkable instruction informs the jury that in order to find self-defense in a murder trial they must find that "The defendant reasonably believed that (he/she/ [or] someone else) was in imminent danger of being killed or suffering great bodily injury…"[170] The obvious difficulty in making such a judgment is determining what it means to "reasonably believe" something. Jurors pressing for some clarification are told: "When deciding whether the defendant's beliefs were reasonable, consider all the circumstances as they were known to and appeared to the defendant and consider what a reasonable person in a similar situation with similar knowledge would have believed." Of course, this is not an explanation but simply

[170] *Judicial Council of California Criminal Jury Instructions* (2015 edition), CALCRIM No. 505.

begs the question of what makes a reasonable person who has reasonable beliefs. But there are no further instructions/clarifications/explanations that judges or lawyers are permitted to give to juries. Lawyers argue and try to persuade the jury of what it means for a belief to be reasonable, and the jurors will likely argue amongst themselves, but there is no more formal guidance for the jury.

What I think this example shows is that there is an undeniable philosophical flavor to the many normative judgments that are made on a daily basis across the entire legal system. The jurors are essentially being asked to make a moral/philosophical judgment of what it means to be reasonable.[171] Under Cavell's conception of philosophy and its methods there is nothing remarkable in this. Any competent user of language can determine what reasonable means. Citizens have shared input into what terms like reasonable mean and access to the community's norms. Of course disagreement is possible, and one might not like the conclusions of the other members of one's community, but there is no priority for one citizen over another. In Williamson's technocratic account of philosophy there is no similar justification.

These sorts of philosophical judgments cannot be delegated away to officials because they are heavily context dependent. Every case is new, in part because contexts are unique and every offender is a unique individual, and thus requires

[171] See Raz (2010) for a lengthy argument that legal interpretations are best understood as both backward looking so that there is fidelity to what the law is and forward looking to how justice can be done in present modern contexts. Raz also argues that legal interpretation is not wholly distinct from moral reasoning.

innovative reasoning on what is just. In one context it may be reasonable for one person to fear for their safety, but a different person might not reasonably be afraid in that same context (*e.g.* someone in outstanding physical shape who has received special forces training). The interpretation must begin anew with each case. There is no general rule that the public can mandate officials follow in their interpretations of terms like reasonableness.[172] And so the judgment must be rendered anew directly by the public. In a meaningful way, they are making it up as they go along. But that is okay, since it is the public's conception of justice that the criminal law legitimately renders.

D. <u>Conclusion.</u>

If we are to continue to consider remorse as grounds for limiting criminal punishments then it is true that we ought to reliably distinguish between sincere remorse and performative remorse. Despite the concerns of cognitive scientists there is good reason to believe that lay judges and jurors are able to reliably do that when afforded the opportunity to observe criminal offenders in court. This is due to the fact that concepts like remorse, sincerity, and deception are best understood in the context of the criminal law through their ordinary meaning as interpreted and judged by competent users of the language. Doing so is in the interest of justice.

[172] Raz 2010.

CHAPTER VI: A DIALOGUE

A philosopher, a scientist, and a lawyer walk into a bar…

Philospher: What is justice?

Lawyer: Great question. But before I offer any counsel I should really inform you that my hourly rate is $550 and that nothing I say here today should be taken to formalize an attorney-client relationship between myself and either of you…

Scientist: I conducted a quick survey of human history and justice appears to be something that philosophers talk about but I was unable to find any evidence of it actually existing.

Lawyer: …and you'll have to agree to arbitration in the event that any of the advice I give you here today causes injury; be it physical, emotional, legal…

Philosopher: But, like in an actual case…when is it justified to punish someone for breaking the law?

Scientist: Well, first we'd have to establish that it is objectively true that they broke the law.

Philosopher: But there is no such thing as objective truth.

Scientist: What!?! Of course there is objective truth.

Philosopher: Hate to break it to you but there are only perspectives.

Lawyer: Yeah, she's right. Objective truth is whatever you can get the twelve people in the box to agree to.

Scientist: Okay, then prove to me there is no objective truth.

Philosopher: You mean you want me to show it to be objectively true that there is no objective truth?

Lawyer: She's got you there. It's like when I told my daughter there was no Santa Clause and she told me she wouldn't believe it unless Santa himself admitted it. Smart kid. Destined to make partner.

Scientist: Well then how do I know you're right?

Philosopher: It feels like you just changed wording but it's the same question.

Lawyer: Good strategy. Just keep asking until you get the answer you want.

Philosopher: It's like I invited you to play soccer and I told you the rules but you keep insisting on using your hands. We can play using our hands, but then it's not soccer.

Scientist: I'd never cheat…or p-hack.

Philosopher: My claim above was not offered as a statement of fact. It is a perspective, as all human statements are. And those are the only sort of statements we have any knowledge of. It is not subject to being proven objectively right or wrong; true or false; correct or incorrect. That sort of move is only valid in certain games/contexts. Philosophy is not such a game.

Scientist: Oh so your claim is above critique. And then all claims are equal. There is no right or wrong. Anything goes.

Philosopher: Don't be scared. Obviously, not all claims are equal. We can ask if any particular claim is valuable. Useful. Beautiful. Pleasurable. Apt. Moral. Healthful.

Lawyer: Billable.

Scientist: But then everything would always be up for argument and whoever was more persuasive would prevail even if it was unjust.

Lawyer: Doesn't sound all bad.

Scientist: So it's okay to punish someone even if it's not true that they broke the law?

Philosopher: Well, no, that doesn't seem right. They have to be responsible in some way…agency…causation…

Scientist: But that's an empirical question isn't it? The sort of thing that can be systematically studied, chronicled, measured? The sort of thing we can get right or wrong?

Philosopher: intention…harm…luck…

Lawyer: You're stalling. Answer the question or I'll move to have you held in contempt.

Scientist: Yeah! Accuracy matters. You have to admit it. You can't just make up the facts or history. Just punishment requires you to get things right.

Philosopher: Well I never said there was no room at all for the perspective that there is objective truth! I guess we have to adopt the perspective of objective truth in certain contexts to achieve specific goals. We just can't expect that since it works so well in those instances that it will work well in every case.

Scientist: So what cases is it useful in.

Lawyer: Every case is unique and requires the advice of a highly trained, and well-compensated, specialist.

Philosopher: You haven't been very helpful. Don't you ever have a case that bothers you because it seems unjust? Don't you wonder about why the rules are the way they are? Don't you think?

Lawyer: Don't look at me! I don't have a dog in this fight. I just need to know what the rules are and then I do my thing. I don't have an opinion on what the rules ought to be. I'm more like a mechanic than an engineer.

Philosopher: But aren't you curious?

Lawyer: I only look under the hood when a customer asks me to. And even then I only pull the levers that are useful in that instance.

Scientist: Ignore him. Look, we know that drawing inferences from empirical observation of past events is our best tool for predicting future events. And it has been highly useful in engineering new technologies for specific tasks. So why don't we use it to engineer a system that achieves justice.

Philosopher: Great! So, then, what is justice?

BIBLIOGRAPHY

Cases and Statutes

In re Shelley Luther, No. 20-0363 (Tex. 2021).

Echavarria v. State, 839 P.2d 589 (Nev. 1992).

California Penal Code § 502

United States v. Scheffer, 523 U.S. 303 (1998).

Riggins v. Nevada, 504 U.S. 127 (1992).

United States v. Semrau, 693 F.3d 510, 6th Cir. (2012).

Fed. R. Evid. § 401-403.

Fed. R. Evid. § 804.

Cal. Evid. Code § 140

Fed. Rule of Evid. § 401

Lambert v. California, 355 U.S. 225 (1957).

Judicial Council of California Criminal Jury Instructions (2015 edition), CALCRIM No. 505.

Duncan v. Louisiana, 391 U.S. 145 (1968)

United States v. Peterson, 483 F.2d 1222, Ct. of Appeals, District of Columbia (1973).

Model Penal Code § 3.04

People v. Goetz, 68 N.Y.2d 96, NY Ct. of Appeals (1986).

18 U.S.C. § 922(g)(1)

California Penal Code § 120022

New Jersey v. Brown, 143 N.J. Super. 571 (1976).

United States v. Dougherty, 473 F.2d 1113 Ct. of Appeals, District of Columbia (1972).

Other References

Alexander, Michelle. The New Jim Crow, (The New Press 2012).

Aristotle. Nicomachean Ethics. Terence Irwin, transl. Second Edition, (Hackett Publishing Company, Inc. 1999)

Austin, J.L. "A Plea for Excuses: The Presidential Address," *Proceedings of the Aristotelian Society,* New Series, Vol. 57 pp. 1-30 (1956-57).

Bagaric, M. & Amarasekara, K. "Feeling Sorry? Tell Someone Who Cares: The Irrelevance of Remorse in Sentencing," *The Howard Journal of Criminal Justice,* Vol. 40(4), pp. 364-376 (2001).

Barrett, Lisa Feldman. "The Law's Emotion Problem," *New York Times*. (March 11, 2017.)

Barrett, Lisa Feldman; Adochs, R.; Marsella, S.; Martinez, A.M.; & Pollak, S.D. "Emotional Expressions Reconsidered: Challenges to Inferring Emotion from Human Facial Movements," *Psychological Science in the Public Interest* Vol. 20(1) 1-68 (2019).

Baron-Cohen, S, & Cross, P, "Reading the eyes: evidence for the role of perception in the development of a theory of mind," *Mind and Language,* 6, 173-186 (1992).

Baron-Cohen, S.; Campbell, R.; Karmiloff-Smith; A., Grant, J.; & Walker, J. "Are children with autism blind to the mentalistic significance of the eyes? *British Journal of Developmental Psychology*, 13, 379-398 (1995).

Baron-Cohen, S.; Wheelwright, S.; & Jolliffe, T. "Is there a "Language of the Eyes"? Evidence from normal adults and adults with autism or Asperger Syndrome," *Visual Cognition*, 4, 311-331(1997).

Baron-Cohen, S.; Wheelwright, S.; Hill, J.; Raste, Y.; & Plumb, I. "The 'Reading the Mind in the Eyes' test revised version: a study with normal adults, and adults with Asperger Syndrome or high-functioning autism," *Journal of Child Psychiatry and Psychiatry,* 42, 241-252 (2001).

Baron-Cohen, Simon. The Essential Difference, (Basic Books 2004).

Basu, Rima. "The Wrongs of Racist Beliefs," *Philosophical Studies* 176(9): 2497-2515 (2019).

Caldwell, Peter C. "Controversies Over Carl Schmitt: A Review of Recent Literature," *The Journal of Modern History* Vol. 77, No. 2, pp. 357-387 (June 2005).

Cavell, Stanley. Must We Mean What We Say? (Cambridge University Press, 1958).

Cavell, Stanley. The Claim of Reason. (Oxford University Press, 1979).

Deigh, John. "The Sources of Moral Agency: *Essays in Moral Psychology and Freudian Theory*," pp. 48-52 (Cambridge University Press, 1996).

Devlin, Lord Patrick. "Morals and the Criminal Law," The Enforcement of Morals, (Oxford University Press, 1965).

Dworkin, Gerald. "Paternalism" from Morality and the Law, ed. R. Wasserstrom (Belmont, CA: Wadsworth Publishing, 1971).

Dworkin, G. & Blumenfeld, G. "Punishments for Intentions," *Mind* New Series Vol. 75 No. 299, pp. 396-404 (July 1966).

Dworkin, Ronald. Taking Rights Seriously, (Harvard University Press, 1977).

Dworkin, Ronald. Law's Empire, (Fontana Press, 1986).

Edmundson, William. "Is Law Coercive?" *Legal Theory* 1: 81-111 (1995).

Feinberg, Joel. The Moral Limits of the Criminal Law: Harm to Others, (Oxford University Press 1984).

Fine, Cordelia. Delusions of Gender, (W.W. Norton & Company, Inc. 2011).

Foucault, Michel. Discipline and Punish: The Birth of the Prison (Alan Sheridan, transl. Vintage Books Edition, 1995).

Goldstein, Abraham. "Conspiracy to Defraud the United States," 68 Yale L.J. 405 (1959).

Hage Johnson, Carla Ann. "Entitled to Clemency: Mercy in the Criminal Law," *Law and Philosophy*, vol. 10, pp. 109-118 (1991).

Haney, Craig. "Counting Casualties in the War on Prisoners," *University of San Francisco Law Review*, vol. 43, pp. 87-138 (2008).

Haney, Craig and Philip Zimbardo. "The Past and Future of U.S. Prison Policy: Twenty-five years after the Stanford Prison Experiment," *American Psychologist* pp. 709-727 (1998).

Hart, H.L.A. "A Prolegomenon the Principles of Punishment," in *Punishment and Responsibility: Essays in the Philosophy of Law* (Oxford: Clarendon, 1968).

Hart, H.L.A. "Immorality and Treason," Philosophy of Law, 9th ed. Edited by J. Feinberg, J. Coleman, and C. Kutz (Wadsworth, Cengage Learning, 2014).

Hayne, Robert Y. "Shall the Jury System be Abolished," in *The North American Review* Vol 139 No. 335 pp. 348-355 (Oct., 1884).

Hobbes, Thomas. Leviathan. Edited by Edwin Curley (Hackett Publishing, Inc., 1994).

Janaway, Christopher. "Autonomy, Affect, and the Self in Nietzsche's Project of Genealogy," Nietzsche on Freedom and Autonomy, ed. By Gemes, K. and May, S. (Oxford University Press 2009).

Kadish, Sanford and Stephen J. Schulhofer. Criminal Law and its Processes (Aspen Publishers 2001).

Kutz, Christopher. "Responsibility," Oxford handbook of Jurisprudence and Philosophy of Law, ed. By Coleman, J. Shapiro, S. & Himma, K.E. pp. 548-587 (Oxford University Press, 2004)

Michel, Lou and Dan Herbeck. American Terrorist: Timothy McVeigh and the Oklahoma City Bombing (Harper Collins 2001).

Mill, John Stuart. On Liberty, (Dover Thrift Edition 2002).

Mills, Charles W. The Racial Contract, (Cornell University Press 1999).

Monk, Ray. How to Read Wittgenstein. (W. W. Norton 2005).

Moore, Michael S. "The Moral Worth of Retribution," Responsibility, Character, and the Emotions, ed. By Schoeman, F. (Cambridge University Press 1988).

Morris, Herbert. Guilt and Shame, (Wadsworth Publishing, Inc. 1971).

Morris, Herbert. "Persons and Punishment," On Guilt and Innocence: Essays in Legal Philosophy and Moral Psychology, (University of California Press 1976).

Murphy, Cullen. God's Jury: The Inquisition and the Making of the Modern World, (Mariner Books 2013).

Murphy, Jeffrie G. Getting Even: Forgiveness and its Limits (Oxford University Press 2003).

Murphy, Jeffrie G. "Mercy and Legal Justice," in J. Murphy & J. Hampton, Forgiveness and Mercy, (Cambridge: Cambridge University Press, 1988).

Nietzsche, Friedrich. "Thus Spoke Zarathustra," The Portable Nietzsche, transl. Walter Kaufmann (Viking Press 1954)(a).

Nietzsche, Friedrich. "Twilight of the Idols," The Portable Nietzsche, transl. Walter Kaufmann (Viking Press 1954)(b).

Nietzsche, Friedrich. Beyond Good and Evil, Walter Kaufmann (Vintage Books 1966).

Nietzsche, Friedrich. On the Genealogy of Morals, Walter Kaufmann (Vintage Books 1969).

Nietzsche, Friedrich. The Gay Science. transl. Walter Kaufmann (Vintage Books 1974).

Plato. "Crito," in Five Dialogues, 2nd ed. Transl. by G.M.A. Grube, Rev'd. by John M. Cooper (Hackett Publishing, Inc. 2002).

Rawls, John. A Theory of Justice, (Belknap Press 2nd ed., 1999).

Raz, Joseph. "On the Nature of Law," *Archive fur Rechts und Sozialphilosophie*, Vol. 82, pp. 1-25, (1996)(a).

Raz, Joseph. "Why Interpret?" *Ratio Juris*, Vol. 9, pp. 349-363, (1996)(b).

Raz, Joseph. Between Authority and Interpretation: On the Theory of Law and Practical Reason, (Oxford University Press 2010).

Richell, R A.; Mitchell, D.G.V.; Newman, C.; Leonard, A.; Baron-Cohen, S.; & Blair, J. "Theory of mind and psychopathy: can psychopathic individuals read the "language of the eyes"?" *Neuropsychologia*, 41, 523-526 (2003).

Rousseau, Jean-Jacques. “Discourse on the Origin and the Foundations of Inequality Among Men,” (Second Discourse) The Discourses and Other Early Political Writings, transl. and ed. by Victor Gourevitch (Cambridge University Press 2000).

Ryle, Gilbert. “Knowing How and Knowing That: The Presidential Address,” *Proceedings of the Aristotelian Society,* New Series, Vol. 46 pp. 1-16 (1945-46).

Sandel, Michael J. Justice: What’s The Right Thing To Do, (Farrar, Straus, and Giroux 2009).

Schmitt, Carl. Political Theology: Four Chapters on the Concept of Sovereignty, transl. by George Schwab (University of Chicago Press 2006).

Shelby, Tommie. Dark Ghettos: Injustice, Dissent, and Reform, (Belknap Press, 2016).

Stanley, Maclen. “Abolish the Jury,” *Psychology Today,* (https://www.psychologytoday.com/us/blog/making-sense-chaos/202111/abolish-the-jury) last accessed (February 16, 2022).

Strawson, Peter. “Freedom and Resentment,” in *FREE WILL* pp. 59-80 (Gary Watson, ed., Oxford: Oxford University Press, 1982).

Ter Hark, Michel. “Uncertainty, Vagueness and Pschological Indeterminacy,” *Synthese,* Vol. 124, No. 2, pp. 193-220 (Aug., 2000).

Vinx, Lars. “Carl Schmitt,” *The Stanford Encyclopedia of Philosophy* URL: https://plato.stanford.edu/archives/fall2019/entries/Schmitt (Fall 2019).

Waldron, Jeremy. “A Right to Do Wrong,” *Ethics,* Vol. 92, No. 1, Special Issue on Rights, pp. 21-39 (Oct., 1981).

Westen, Peter. “Why the Paradox of Blackmail is so Hard to Resolve,” in *Ohio State Journal of Criminal Law,* Vol. 9:2 pp. 585-636 (2012).

Whitehurst, Lindsay and Michael Tarm. “Ruling raises new questions about remote testimony in court: An overturned conviction in Missouri is raising new questions about video testimony in criminal court cases nationwide,” (Associated Press January 13, 2022) *last accessed at* https://abcnews.go.com/Politics/wireStory/ruling-raises-questions-remote-testimony-court-82249190 (on February 17, 2022).

Williams, Bernard. "Nietzsche's Minimalist Moral Psychology," Nietzsche, Genealogy, Morality. Ed. Schacht, R. (Berkeley: University of California Press 1994).

Williams, Glanville. Criminal Law: The General Part 2 (2nd ed. Stevens and Sons, Ltd. 1961).

Williamson, Timothy. The Philosophy of Philosophy. (Oxford: Blackwell Publishing 2007).

Wittgenstein, Ludwig. Remarks on the Philosophy of Psychology Vol. I, (G.E.M Anscombe transl., Basil Blackwell Oxford) 1980.

Wittgenstein, Ludwig. Remarks on the Philosophy of Psychology Vol. II, (C.G. Luckhardt and M.A.E. Aue transl., University of Chicago Press) 1980.

Wittgenstein, Ludwig. Last Writings on the Philosophy of Psychology Vol. I, (C.G. Luckhardt and M.A.E. Aue transl., Basil Blackwell Oxford) 1982.

Wittgenstein, Ludwig. Last Writings on the Philosophy of Psychology Vol. II, (C.G. Luckhardt and M.A.E. Aue transl., Basil Blackwell Oxford) 1992.

Wittgenstein, Ludwig. Philosophical Investigations, Revised 4th Edition (G.E.M. Anscombe, P.M.S. Hacker and Joachim Shulte transl., Wiley Blackwell) 2009.