

**Reclaiming responsibility: an ameliorative proposal*****Reivindicando responsabilidades: uma proposta de melhoria*****Michelle Cieurria¹**

University of Missouri-St. Louis

ABSTRACT

This paper considers and critiques two models of moral responsibility: the retributive model, which sees holding-responsible as a retributive or punitive response to a wrongdoer, and the rehabilitative model, which sees holding-responsible as a tool for agency enhancement, moral improvement, or moral scaffolding. I argue that, while responsibility can serve these purposes, it should do so in a way that (also) addresses oppression. I call this view “ameliorative” because it seeks to ameliorate the responsibility system through social justice. In making this argument, I underscore the connections between retribution and rehabilitation *in practice* (i.e., in the real world) and structural injustice. These connections are too often elided or overlooked in optimistic accounts of responsibility’s retributive and rehabilitative potential, in contrast to their bleak reality.

Key-words: moral responsibility; retribution; rehabilitation; ideal theory; ameliorative.

RESUMO

Este artigo considera e critica dois modelos de responsabilidade moral: o modelo retributivo, que vê a responsabilização como uma resposta retributiva ou punitiva a um transgressor, e o modelo reabilitativo, que vê a responsabilização como uma ferramenta para o aprimoramento da agência, melhoria moral, ou andaimes morais. Defendo que, embora a responsabilidade possa servir estes propósitos, deve fazê-lo de uma forma que (também) aborde a opressão. Chamo esta visão de “melhoradora” porque procura melhorar o sistema de responsabilidade através da justiça social. Ao apresentar este argumento, sublinho as ligações entre retribuição e reabilitação na prática (ou seja, no mundo real) e a injustiça estrutural. Estas ligações são muitas vezes omitidas ou negligenciadas nas descrições otimistas do potencial retributivo e reabilitador da responsabilidade, em contraste com a sua triste realidade.

Palavras-chave: responsabilidade moral. retribuição. reabilitação. teoria ideal. melhoramento.

¹ E-mail: Mich.cieurria@gmail.com. ORCID: <https://orcid.org/0000-0002-8722-1003>.



INTRODUÇÃO

In this paper, I will consider and critique two popular views of responsibility. I call the first view *retributivism*. On this view, holding someone responsible involves expressing negative attitudes (like anger and resentment) toward the person and judging those attitudes to be deserved or fair. Retributivists typically want the target to feel the sting of moral judgment, penalties, or sanctions. I call the second view *rehabilitationism*. On this view, holding someone responsible involves judging the person to be an apt candidate for moral improvement, enhancement, education, or scaffolding. Rehabilitationists want to enhance the target's agency.

Although these may not be the labels typically applied to these views, I have chosen them to highlight the relationship between theoretical constructs of responsibility and real-world practices of retribution and rehabilitation - practices critiqued by many feminist, crip, queer, and decolonial theorists. While the aspirations of righteously punishing wrongdoers and benevolently enhancing wrongdoers' agency may seem noble, the reality of unjust punishment and rehabilitation - injustices that are structural and systemic - should give us pause.

In this paper, I will apply what Charles Mills (2017) called "non-ideal" analysis to underscore the risks associated with deploying retributive and rehabilitative responsibility in structurally unjust contexts - contexts in which retributive and rehabilitative systems are cruel and exploitative. Non-ideal theory focuses on the existence of injustice, oppression, and inequality in the real world, rather than imagining an ideal world in which these issues do not arise, or are infrequent and accidental. My aim is not to undermine the viability of retributive and rehabilitative responsibility (as 'eliminativists' seek to do [e.g., Levy, 2011; Waller, 2014]), but rather to emphasize that when we use these forms of responsibility, we must be sensitive to the structural injustices that shape commonsense intuitions and institutional norms. A politically nuanced and historically informed approach is needed to correct inequalities in socially-embedded responsibility praxis and to build a better responsibility system.

With this in mind, I propose an alternative account of responsibility that views retributive and rehabilitative uses of responsibility as lower-order functions, which ought to be governed by a higher aim: that of ameliorating responsibility

praxis by addressing oppression and injustice in the norms and expectations around responsibility. I call this view “ameliorative” based on Sally Haslanger’s (2012) model of ameliorative analysis as a philosophical method that seeks to not only analyze and describe everyday concepts but also to improve them.

In what follows, I will explain and critique retributive and rehabilitative accounts of responsibility, laying the groundwork for an ameliorative model. I will then offer examples of how I think the ameliorative model should work in practice, focusing on the use of responsibility within the disability justice movement, to which I belong as a disabled, queer, gender-fluid activist. My “criptiques” (Wood, 2014) of standard responsibility praxis are influenced by my first-hand knowledge of how retributive practices affect queer, neurodivergent, disabled folks through invasive systems of policing, incarceration, and self-discipline, and how rehabilitative practices operate through coercive systems of pathologization, medicalization, and forced treatment. My aim is not to provide a comprehensive analysis but to illuminate these undertheorized aspects of responsibility praxis.

1. RETRIBUTIVE RESPONSIBILITY

The classic and perhaps best-known approach to responsibility is (what I will call) the retributive view. The retributive view says that responsibility functions or should function to punish wrongdoers for inexcusable wrongdoing. In *The Stubborn System of Moral Responsibility* (2015), Bruce Waller cites Michael S. Moore (2010), Peter French (2001), and Robert C Solomon (1990) as retributivists. French writes, “wrongful actions require hostile retribution (2001, p. 97).” Solomon says, “to seek vengeance for a grievous wrong, to revenge oneself against evil—that seems to lie at the very foundation of our sense of justice, indeed, of our very sense of ourselves, our dignity, and our sense of right and wrong” (1990,: 37) Similarly, “Susan Wolf (2011), David Shoemaker (2015, 2017), and Leonhard Menges (2017)... defend accounts of blame that emphasize anger” in response to wrongdoing (Tognazzini & Coates 2024). Strawsonians, too, tend to believe that “negative reactive attitudes” like resentment and anger are fitting responses to wrongdoing (Strawson, 2008). One justification for these negative attitudes is retribution. As Gregg Caruso and Derk Pereboom put it,

When we confront a wrongdoer with expressions of angry blame (Wolf 2011; Fricker 2016; Bagley 2017; Shoemaker 2018; McKenna 2019), such blame is associated with a certain negative emotional attitude toward the wrongdoer - such as resentment or indignation; and more broadly, moral anger - that goes beyond the mere absence or withdrawal of good will (Wolf, 2011, :p. 335). This kind of moral blame is non-trivially painful or harmful, and, when accompanied with a supposition of desert, qualifies as *retributive*.

Not everyone believes that blame should be retributive, however. Critiques come from other camps, including eliminativism (Waller's view) and rehabilitativism, as well as marginalized discourses. In what follows, I will outline some of these critiques to underscore the limitations of the retributive model, making space for an alternative. I will not, however, try to refute retributivism, or argue that blame should never be retributive in nature. Instead, I will argue that retributive blame should be guided by a higher aim - namely, an ameliorative aim. That is, blame should ameliorate responsibility praxis by mitigating oppression and injustice. To simplify things, I will focus on two influential accounts of retributive responsibility.

Michael McKenna (2012) defends a version of retributivism (in Caruso and Pereboom's sense). He argues that moral responsibility is akin to a conversational exchange with three stages: (1) S makes a moral contribution in the form of a transgression, (2) P morally addresses S or asks for a response, and (3) S offers P an account of their behavior in the form of an acknowledgement, apology, or other form of redress. McKenna's view is a form of Strawsonianism in the sense that moral conversation is taken to involve an exchange of reactive attitudes such as anger, resentment, and indignation, and these attitudes are taken to be deserved responses to a person's negative intentions, attitudes, or "quality of will."

Next, let's turn to Susan Wolf's account of angry blame. Wolf says that "angry blame," which encompasses such emotions as resentment, indignation, and righteous anger, is a justified response to people who have violated personal relationships. When you angrily blame someone,

you may want to 'get in his face.' You may want him to see your anger and to feel your pain. You may want to scold him; you probably want an apology. And although the angry emotions and attitudes do seem to me to be conceptually tied to a disposition to punish, and therefore with a willingness to make the object of blame suffer in a particular way, it would be a serious mistake to identify this with a general withdrawal of good will (Wolf, 2011, p. 12).

Angry blame motivates you to confront the wrongdoer. Hence, it “may bring people together and make them closer, rather than pushing them away” (ibid.). Wolf thinks that moral repair is especially likely in the context of “family life,” and uses the example of a hypothetical Italian family whose anger is constructive. Yet she acknowledges that retribution has limits:

It may be noted, however - perhaps with impatience - that all the examples I have used to make my case for angry blame concern relatively trivial acts by reasonably decent people in the context of fairly intimate relationships. (Welcome to my world!) Even if the liability to angry blame may appear to have an acceptable place in these contexts, we all know how dangerous such blame can be in other circumstances (2011, p. 12).

One such circumstance is the prison system:

One proposal worth considering would restrict the appropriateness of angry blame, or at least of its physical expression, to private contexts of interpersonal relationships, and advocate that, in public contexts, and especially in connection with the criminal justice system, we assess and revise our practices exclusively with the more dispassionate model of [dispassionate] Scanlonian blame in mind (2011, p. 15).

For “criminal wrongdoing,” a restorative justice model may be more appropriate.

In drawing this distinction, Wolf is invoking the popular critique from rehabilitationists and eliminativists that retributive blame - perhaps especially angry and hostile forms thereof - is inseparable from the violence of over-incarceration, racial incarceration, unpaid penal labor, and other forms of modern colonialism. Waller, in defense of eliminativism, says that our retributive impulses motivate “tougher measures against criminals (such as mandatory sentencing, minimum sentences, ‘boot camps’ and adult prisons for juvenile offenders, ‘sex offender’ permanent registrations, and ‘three strikes’ laws)” (2015, p. 50). Thus, he believes that we should suppress or expunge our retributive impulses, not only in ‘public settings’ but in general. He is against the use of ‘strike-back’ blame in any context. This is, in part, because agrees with Wolf that blame is “conceptually tied” to punishment and suffering, but he takes this to be a reason to abolish blame entirely.

Contrary to Waller’s critique, some feminists have defended the value of anger and retributive blame. Audrey Lorde says, “my anger is a response to racist at-

titudes and to the actions and presumptions that arise out of those attitudes” (1984, p. 124). For Lorde, anger can be a source of clarity, a catalyst for change, a source of connection, and a means of survival. Perhaps angry blame, too, can serve these purposes. Macalester Bell defends the value of “hostile blame” as a vehicle for educating people, motivating people to do better, and standing up to oppression. Bell believes that retributive reactions can be salutary and socially beneficial (2012, 2013). This does not, however, negate the worry that retributive blame can be, and often is, deployed in oppressive ways. Perhaps the problem is not retributive blame *per se*, but rather the way retributive blame is expressed, especially towards members of oppressed groups. Maybe blame should be guided by a higher aim, such as addressing oppression.

This is precisely what intersectional feminist critiques suggest. Neither angry blame nor retribution is inherently harmful, but the way these practices are performed in our society is often oppressive. Marilyn Frye (1983) defines “oppression” as a system of interrelated barriers and forces that reduce, immobilize, and shape people, particularly minorities, forcing them to live confined and restricted lives, as if in a bird cage. One of the vehicles for oppression, according to Frye, is the “double-bind,” which creates lose-lose situations for members of oppressed groups. For instance, if a woman is sexually active, she is blamed for being “loose, unprincipled, or a whore,”; but if she “refrains from heterosexual activity,” she is blamed for being “frigid,” “upright,” “a bitch,” or a “dike” (1983, p. 3). This example illustrates how blame, along with other punitive practices, can impose double-binds on oppressed minorities, upholding systems of oppression such as patriarchy, cissexism, and colonialism.

Jules Holroyd (2021, 2024) emphasizes that praise, which is often thought of as anodyne, can be just as oppressive as blame. For instance, disabled people are often praised for ‘overcoming’ their disability or engaging in everyday activities like attending work or school, implying that disability is nothing more or less than an impairment or tragic flaw. This type of praise, often labeled ‘inspiration porn,’ objectifies disabled people as motivational resources for others, and reinforces ableist narratives about disability (as tragic, pathological). On the other hand, disabled people who do not achieve ‘inspirational’ goals may be blamed for lacking ‘resilience’ or ‘optimism,’ creating a ‘supercrip/sadcrip’ double-bind.

Forgiveness, which plays a role in excusing wrongdoing and exempting people from blame, is also asymmetrically distributed. Kate Manne (2017) coined the term “himpathy” to describe the disproportionate sympathy accorded to (privileged) men in patriarchal contexts. Himpathy incites forgiveness, compassion, and a willingness to excuse wrongdoing. The phrase “boys will be boys” is an example of himpathy’s role in patriarchal forgiveness patterns. *He sexually harassed his classmate? Boys will be boys.*

Along similar lines, Shellbi N. Meissner and Andrew M. Smith explain that white settlers benefit from a mystique of innocence created by the “limited histories and futurities” that frame colonial wrongdoing as “a temporal punctuation outside of the norm,” rather than an ongoing pattern of state-sanctioned violence (2024, p. 11). These revisionist histories generate misplaced sympathize for white settlers’ claims to moral purity and goodwill. Furthermore, they contribute to a culture of “structural gaslighting” that “functions to obscure the nonaccidental connections” between “structures of oppression,” colonial harms, and white culpability (Berenstain, 2020). Colonial gaslighting scripts protect white settlers from blame.

These feminist critiques shed light on the background conditions of ‘conversational blame.’ What counts as a transgression - or, as McKenna would put it, what determines the “agent meaning” of a transgression (2012, p. 92) - is influenced by structural injustices such as patriarchy, heteronormativity, and colonialism. Having sex may be socially defined as a transgression for a woman but not a man. Abstaining from sex may also be defined as a transgression for a woman - hence the double-bind. Oppression ensures that some moral discussants have less freedom (from suspicion, blame, retribution, and hostility) than others.

Similarly, oppression ensures that “moral accounts” will not be fairly adjudicated or taken up. Himpathy will incline discussants to forgive privileged men more readily than others for (especially sexist) transgressions. The mystique of settler innocence will give white discussants a presumption of innocence in conversations about responsibility for colonial violence. Gaslighting narratives will make oppressed discussants look guiltier than privileged ones.

What about retributive blame within families? Is it safer than retributive blame in “public contexts”? Some feminist critics argue the opposite. Bell hooks des-

cribes the family as an “institutionalized sphere of power that can easily be autocratic and fascistic” (2000, p. 20) In the context of what hooks calls “white supremacist capitalist patriarchy” (1997), oppressed family members, especially children, are highly susceptible to abuse. (In the United States, children witness and experience more violence than adults, yet are the most politically disenfranchised social group, with no right to vote, participate in politics, or earn money)². Similarly, Sophie Lewis notes that “the family is where most of the rape happens on this earth, and most of the murder” (2002, p. 62). Lewis argues that the family is “the name we use for the fact that care is privatized in our society,” an arrangement that leaves “the ill, the young, and the elderly” highly susceptible to domestic violence, neglect, and abuse (2002, p. 8). Blame plays a significant role in these family dynamics. Children, elders, and disabled people are highly susceptible to angry, hostile, and abusive blame in ‘autocratic and fascistic’ families. The rate of unjust angry blame may, in fact, be higher in private than public contexts, because of the privacy one has within the home compared to public settings. This risk is compounded by the depoliticization of domestic life by classic social contract theorists like John Rawls, who (mis)represented the home as an apolitical sphere of voluntary caregiving. In *The Sexual Contract* (1988), Carole Pateman argues that domestic abuse is masked by the pretext that the home is a ‘private sphere’ of intimate relationships, separate from the messy world and often violence world of politics. In reality, the home may be even more violent than public spaces, especially given the patriarchal nature of the traditional family.

When it comes to “criminal wrongdoing,” Wolf suggests that we should suspend angry blame and adopt a restorative justice approach, as angry blame could exacerbate carceral violence. But is the problem that prisons use *excessively retributive* and *angry* forms of responsibility-holding, or that they don’t hold people responsible at all – that is, responsibility-holding isn’t one of the functions of imprisonment? According to Angela Davis (2011), the main purpose of the prison-industrial complex is not to hold people responsible for crimes, or to punish them, or to rehabilitate them. These reasons are all smoke and mirrors. The true purpose of imprisonment is to generate profits for corporations; “To understand the social meaning of the prison today within the context of a developing prison industrial com-

² See the National Youth Rights Association website for more information on the importance of youth rights and youth suffrage: <https://www.youthrights.org/about/what-are-youth-rights/>

plex,” says Davis, “means that punishment has to be conceptually severed from its seemingly indissoluble link with crime” (2011, p. 85). Punishment is for profit, *not* for the purpose of holding criminals responsible, rehabilitating them, or reducing the crime rate - in fact, America’s Global Domestic Product depends on *increasing* the prison population. Thus, it is not clear that retributive responsibility is a significant factor in carceral injustices such as overincarceration and racial incarnation, since it is not clear that responsibility is a causative reason for incarceration. What *is* a reason, says Davis, is private profits, along with racism, sexism, ableism, and other oppressions that make certain groups especially prone to criminalization. All of these oppressions are inextricable linked with a capitalist profit motive.

None of this refutes the value of retributive blame, but it suggests that it is often, and perhaps systematically, expressed, deployed, and distributed in oppressive ways within unjust societies. Nonetheless, feminist critiques by Audre Lorde, Macalester Bell, and others suggest that retributive blame *could* be beneficial if deployed correctly, i.e., consistent with principles of justice. Perhaps, then, retributive blame should not aim solely at retribution, but should serve a higher purpose or goal. I suggest that this higher-order goal is to address, mitigate, and call out oppression. If retributive blame is used at all, it should be used to dismantle “the master’s house,” in Lorde’s terms (2012), not *merely* to punish wrongdoers. This use of retributive blame can, I think, be seen in *The Uses of Anger*, in which Lorde angrily held white women responsible for denying their own complicity in racist systems. While Lorde does not endorse harsh punishment, she does believe that people should face the consequences of their actions, including penalties and sanctions that fall under the remit of retribution. Thus, angrily calling out racism could be an ameliorative use of retributive blame.³

2. REHABILITATIVE RESPONSIBILITY

Another popular theory of responsibility is what I will call the rehabilitative model. Under this label, I include agency-enhancement views, moral-improvement views, and moral-scaffolding views. One of the most popular versions of this theory

³ This may be a controversial reading of Lorde, but I believe that one can reasonably interpret her as supporting certain retributive practices, as long as they respect the target and do not impose undue hardships.

if Manuel Vargas' agency-enhancement view. This view (in its original formulation) says that responsibility can help us "build better beings" by encouraging actions that are conducive to human flourishing and social cooperation (2013). Vargas believes that we can use responsibility to shape and enhance individuals' capacities for self-regulation and moral agency, teach people about societal norms and values, encourage people to reflect on their actions and intentions, and foster a sense of personal responsibility, making people more reliable and cooperative moral agents.

Another example of this view is advanced by Gregg Caruso and Derk Pereboom (2022). They reject retributivism and advocate for a forward-looking account of responsibility that promotes moral education, moral improvement and rehabilitation. Caruso in particular advocates for a "public health-quarantine" model, which treats harmful behavior as a public health issue rather than a matter of moral fault or guilt. This approach emphasizes harm reduction through social interventions, education, and therapeutic measures. Caruso specifically uses the language of "rehabilitation" to describe the aims of this model. He and Pereboom believe that the public-health model is especially suitable for "dealing with dangerous criminals" (2022, p. 55), who should be "quarantined" and rehabilitated rather than punished.

The rehabilitative approach in general is often taken to be a gentler, kinder, and more "civilized" (to use Strawson's language [2008, p. 10]) alternative to the retributive model. Rather than punishing wrongdoers, wouldn't it be more humane to quarantine, treat, and rehabilitate them?

Feminist philosophers tend to agree that rehabilitation and moral repair are valuable goals. Margaret Urban Walker believes that moral responsibility is not only about individual accountability, but also about how individuals and communities engage in the broader process of repair and restoration. She defines "moral repair" as "restoring or creating trust and hope in a shared sense of value and responsibility" (2006, p. 28). Patricia Hill Collins similarly believes that reconciliation and repair help to address historical injustices, restore dignity, build trust, promote social justice, and encourage collective responsibility (2022). On both views, moral responsibility can facilitate rehabilitation and moral repair, and vice versa.

Having said this, feminists also draw attention to the dark side of rehabilitation, or at least what often passes for 'rehabilitation' in unjust societies. This is an important critique, since ignorance of the history and politics of rehabilitative pra-

xis can lead to confusion about what constitutes genuinely rehabilitation versus harmful facsimiles thereof - the type that have been institutionalized systematically throughout the history of America. I will refer to these facsimiles as ‘rehabilitation*’ to distinguish them from their genuinely restorative counterparts. Arguably, rehabilitation* is more prevalent than genuinely reparative rehabilitation in mainstream society due to ongoing forms of oppression and state-sanctioned violence.

Rehabilitation* has a long history and can still be seen in social institutions ranging from hospitals to prisons to schools to the family. When we examine the history of so-called rehabilitation, we find many examples of abuse advertised as ‘scientific’ rehabilitative practice.

For instance, during the transatlantic slave trade, “Black people kidnapped from Africa and enslaved in the Americas were declared defective as a way to justify and strengthen the institution of slavery,” and as an excuse to “rehabilitate” enslaved people with hard labor, writes Eli Clare (2017, p. 23). The idea was that racialized minorities could be ‘civilized’ or ‘rehabilitated’ into a civilized way of life through slave labor. Dr. Samuel A. Cartwright invented the diagnostic terms “drapetomania” and “dysaesthesia aethiopica” to describe the “pathologies” of wanting to flee from slavery and disobeying white authorities, respectively (ibid.). The cure, according to Cartwright, was hard labor, whipping, and other punishments.

These rehabilitative* practices continued throughout the 20th Century, when civil rights activists, particularly Black Americans, began to be diagnosed with “protest psychosis,” the modern analogue of drapetomania (Metzl, 2010). At the same time, the label of schizophrenia, which had previously been considered an affliction of middle-class white women, was rapidly shifted onto Black men, particularly Black activists like Malcolm X. This diagnosis evolved to include “hostile,” “paranoid,” and “criminal” behaviors (ibid.). This new psychiatric practice of pathologizing Black activists was ‘rehabilitative’ in name only.

Around the same time that civil rights leaders were being pathologized in high numbers, prisons were being reframed as rehabilitative institutions or ‘correctional facilities’ rather than institutions of retribution and punishment. Andrea Pitts notes that throughout the 20th Century,

Prisons and jails began to be considered for their “rehabilitative” effects, rather than serving simply as a means of punishment or deterrence. “Black Codes” in the USA, for example, penalized recently emancipated Black Americans for vagrancy and labor contract violations (e.g., refusing to work in accordance with the terms of a white employer)... The rationale was that prisons and jails, alongside asylums, residential schools, and hospitals, could be used to socially, physically, and emotionally shape those who inhabited these institutions and make them ready to enter the labor force upon release. Such a framing of “rehabilitation” also justified the convict leasing system throughout the US South, a series of programs that treated incarcerated peoples as slaves, and leased incarcerated workers to private companies for pay (2022, p. 240).

‘Rehabilitation’ was quickly replacing punishment as the rationale for mass incarceration, racial incarceration, and forced penal labor that benefited white-run corporations. That is, the ‘public health model’ of criminal justice was *already* in vogue 100 years ago. But the rhetoric of ‘rehabilitation’ and ‘health’ was simply an excuse for the expansion of the prison-industrial complex, a leading source of profits for capitalist elites after the abolition of slavery. ‘Rehabilitation’ was marketed as being in the criminal’s best interests, even though prisons continued to use prisoners as an expendable source of cheap labor.

The logic of rehabilitation was gaining traction in many other spaces as well. Freudians were diagnosing gender-nonconforming women with “wandering uterus,” “hysteria,” and various mood disorders. In *The Feminine Mystique* (1963), Marylin Friedan reported that many white, middle-class, suburban women were being pathologized and sedated by psychiatrists in response to their complaints about domestic servitude: “Many suburban housewives were taking tranquilizers like cough drops,” writes Freidan (1963, p. 20). Psychiatrists were trying to ‘rehabilitate’ stay-at-home moms rather than listening to their political grievances. Around the same time, queer people were being diagnosed with “homosexuality” - a form of ‘sociopathic personality disturbance’ according to the *Diagnostics and Statistical Manual 2nd edition* - and ‘treated’ with conversion therapy, chemical castration, ECT, and sometimes lobotomies. Incarceration was also seen as a potential ‘cure’ for same-sex attraction. These ‘rehabilitations’ had no therapeutic benefit for the patient.⁴

⁴ For a thoughtful examination of addict oppression under the guise of “rehabilitation,” see Virgil Murthy’s “Against arguing about addict agency” (2024).

None of this is to say that rehabilitation is inherently harmful, but rather to underscore that the rhetoric of rehabilitation is often used to justify ‘therapeutic interventions’ that are worse than punishment. The film *One Flew Over the Cuckoo’s Nest* emphasizes that involuntary hospitalization, along with the fashionable treatment regimens of the day (ECT, lobotomy, solitary confinement), can be much worse than a prison sentence (Foreman, 1975). Understanding this history of rehabilitative praxis is important if we want to avoid repeating the injustices of the past - injustices that never went away. Black people are still diagnosed with schizophrenia at higher rates than white people; women are still diagnosed with mood disorders more frequently than men; conversion therapy for ‘homosexuality’ is still practiced in some states; disabled people are still forcibly confined in nursing homes and mental hospitals in violation of the Americans with Disabilities Act (ADA). Rehabilitative praxis is still haunted by the legacy of slavery, heteropatriarchy, and eugenics.

If I am right, then the prison system is already, and has for a long time been, shaped by a rehabilitative* ethos, which has no therapeutic benefit and does nothing to improve prisoners’ lives, foster moral repair, or facilitate social justice. This is because real-world rehabilitative *praxis* is very different than the rosy picture of rehabilitation painted in many theoretical accounts. Often, and perhaps systematically, rehabilitative rhetoric is used to justify forced institutionalization and medical oppression, not to help, heal, or educate oppressed people.

This is one of the reasons for Davis’ insistence that we should be advocating for prison abolition, *not* prison reform, including rehabilitative reforms that have perversely allowed prisons to continue growing (with only a recent marginal decline) under the guise of ‘treatment’ and ‘therapeutical improvement’ for prisoners. This false narrative channels billions of dollars per year into the pockets of private prison CEOs, private contractors that sell services like healthcare, catering, and cleaning to prisons, and corporations that use cheap prison labor to manufacture consumer products like soap, furniture, and clothing (Davis, 2011). A move toward rehabilitative imprisonment will not remedy overincarceration, racial incarceration, or carceral violence, because prisons shifted to a ‘correctional’ model long ago. Instead, we need to move from forced institutionalization toward deinstitutionali-

zation and decarceration. Changing the rationale for imprisonment simply gives prison CEOs a new excuse.

The distinction between rehabilitative theory and rehabilitative praxis is often overlooked and undertheorized in accounts of rehabilitative responsibility. When American bureaucrats have passed bills and policies purporting to *enhance people's agency* or *improve people's morals*, they were, in many cases, using the rhetoric of 'rehabilitation' to oppress and exploit minoritized groups, generating profits for corporate owners and capitalist elites. This is how colonialism works: it mystifies oppression by presenting violence as 'care' and subjugation as 'rehabilitation.' We must therefore regard practices that purport to be rehabilitative with appropriate skepticism, historical sensitivity, and an eye for amelioration.

If rehabilitation is to be socially beneficial, it must be consistent with principles of justice. To me, this suggests that rehabilitation should have a higher aim - that of addressing, mitigating, and reversing oppression. The history of rehabilitation* raises many questions about the direction of responsibility, such as: is it 'criminals' who need to be rehabilitated,⁵ or the owners of corporations that exploit prisoners for profit? Should rehabilitative responsibility be addressed to people who break the law in protest, like Martin Luther King Jr., or to members of the FBI who keep tabs on Black activists? Should disabled people be 'rehabilitated' in nursing homes, or should nursing-home owners be treated for their ableist illusions? Rehabilitative responsibility is only beneficial if it militates against structural injustices.

In light of the history and present context of structural injustice, it is clear that the direction of rehabilitative responsibility must be reversed. If anyone should be rehabilitated through blame and praise, it is the corporate elites who profit from the subjugation of oppressed people.

3. AMELIORATIVE RESPONSIBILITY

In this section, I will introduce a model of responsibility that is compatible with the practices of retribution and rehabilitation but that views these practices

⁵ I put "criminals" in scare quotes because, as Angela Davis affirms, we are all criminals. Who among us has not pirated books, jay walked, or committed some other crime? What determines whether criminals are *criminalized* and *incarcerated* is their skin color, socioeconomic status, and other demographic markers that make certain groups particularly susceptible to policing and sentencing.

as subordinate to a higher purpose: ameliorating the system of responsibility by addressing oppression and structural injustice. Elsewhere (2022, 2020), I have called this model ‘ameliorative’ because it sees the end goal of holding-responsible as ameliorating responsibility or making responsibility praxis less oppressive. This model overlaps with ‘protest accounts’ (Smith, 2013, Hieronymi 2004), but does not limit responsibility to protest alone. Rather, responsibility can do many things - punish, rehabilitate, protest - but whatever it does, it should thereby mitigate injustice. By recognizing that many forms of responsibility, whether retributive or rehabilitative or otherwise, can contribute to an anti-oppressive responsibility system, we expand the range of tools in our activist toolkit. Arguably, pluralistic approaches are the most effective. When Kate Manne discusses the challenges of combatting misogyny, she notes that “its mechanisms and methods are so opportunistic— or enterprising, depending on how you look at it— and so various,” that it can be hard to pin down and effectively address misogyny (2017, p. 30). Oppressions in general tend to shift, evolve, and “capture” (Táíwò, 2022) opposing narratives to evade detection and opposition. Perhaps our response to oppression, then, should be equally diverse, adaptable, and ‘enterprising.’ If retributive, rehabilitative, and protest forms of responsibility can remediate oppression, then any of these strategies will do. We should, in short, use responsibility in as many ameliorative ways as we can.

I borrowed the term “ameliorative” from Sally Haslanger (2000), who advocates for an ameliorative method of philosophical analysis as a tool for social justice. She contrasts this method with traditional “descriptive” and “conceptual” approaches. The descriptive method aims to analyze how a concept is currently understood and used in everyday language and thought, while the conceptual method aims to explore the internal coherence and logical structure of a concept. The ameliorative method goes beyond merely describing and refining existing concepts, focusing on what concepts *should* be, particularly in light of social and political goals. The goal of ameliorative analysis is (paraphrasing Marx) not only to interpret the world in various ways, but to change it. Thus, ameliorative analysis can help us to improve the “world” of responsibility.⁶ When we hold one another

⁶ Lugones (1987) uses the term “world” to refer to set of social norms, practices, values, and power relations that shape how people understand themselves and others.

responsible, we should do so in anti-oppressive ways, thereby ameliorating public understandings and uses of responsibility.

This change we seek, moreover, should be not merely “conservative” or “modestly revisionary,” as some philosophers have argued (Doris, 2015; McGeer, 2015), but positively revolutionary (Tremain, 2024), since only a revolutionary response is sufficient to address structural injustice. Accordingly, we should, as Angela Davis argues, aim not only to reform the prison-industrial complex, but to abolish it. We should, as Sophie Lewis argues, not merely reform the family, but move toward non-privatized, cooperative modes of caregiving. We should, as Shelley Tremain argues (2021), not simply improve nursing homes, but end for-profit institutionalization under the guise of ‘care.’ These critiques show us how ameliorative responsibility should work: it should seek to understand, identify, and dismantle oppressive systems, rather than seeking incremental reform within existing systems of oppression. Neoliberal or reformist policies do not go far enough to address entrenched injustices.

By focusing on the aim of amelioration when practicing responsibility, we shift attention to the “seamy side”⁷ of retributive and rehabilitative practices, especially their ties to historic injustices. While retribution and rehabilitation can *ideally* enhance the social good, genealogical analysis reveals that, in real-world conditions, this is not what these practices typically do. In colonial societies, ‘rehabilitative’ practices tend to support racial incarceration and carceral violence. In patriarchal societies, ‘rehabilitative’ practices pathologize women’s grievances. An ameliorative lens highlights these oppressive dimensions of responsibility praxis, consistent with “non-ideal theory” - a method of analysis defended by Mills as a way of avoiding unrealistic idealizations by focusing attention on real-world injustices, inequalities, and the testimony of oppressed groups. Retributive and rehabilitative responsibility can be beneficial if they are responsive to structural injustices, not if they replicate and reproduce them. It’s up to us to use them in the correct way. This means that we must not only educate ourselves on oppressi-

⁷ Strawson (2008) used the term “seamy side” to describe distortions in our responsibility practice, such as “self-deception” and “guilt-transference,” but he says that these issues are of “minor importance.” In contrast, I believe that they are of major importance, as does Jules Horoyd (2024), who has explored the shortcomings of Strawson’s claim at some length.

on, but elevate the testimony of oppressed groups, raising awareness about the entrenched harms of ordinary responsibility praxis.

So far, I have concentrated on the oppressive dimensions of responsibility - the doom and gloom side of moral life - but I do not want to end on this note. Instead, I will provide examples of how activists are reclaiming responsibility and using it in ameliorative ways. This will provide a model of 'best practice' for holding people, groups, and collectives responsible.

I will focus on a movement with which I am intimately familiar: disability justice. As discussed above, disabled people are objects of 'inspiration porn' when they achieve able-bodied markers of success, such as getting a job or getting married, and are objects of 'pity porn' when they do not. Disabled activists are disrupting this double-bind by promoting realistic narratives about disability, rooted in their lived experiences. For example, the documentary *Crip Camp* (Freeman & LeBrecht, 2020) shines a light on the intergenerational struggles of disabled activists against ableist oppression - something for which activists deserve credit and praise. Yet disabled people are often perceived as incapable, incompetent, and helpless. Indeed, they are even blamed and held in contempt for demanding 'too much' change or 'too many' resources. In *Crip Camp*, Judy Heumann expresses her indignation at being expected to show gratitude for equal access to restrooms. She says, "I'm very tired of being thankful for accessible toilets," which underscores that accessible toilets are a right, not a privilege. Therefore, disabled people are not blameworthy for demanding access to restrooms. Furthermore, they do not owe able-bodied policymakers praise for passing legislation that codifies this right. Providing access to toilet is not a noble or supererogatory act; it is a duty.

To combat these ableist assumptions about who deserves responsibility for what, disabled activists have created the disability pride movement, which celebrates disability and challenges ableist narratives (Mattlin, 2022). Among other things, this movement empowers disabled people to take responsibility for their choices, good and bad, rather than allowing able-bodied society to decide what they are and are not responsible for. Through disability pride, disabled people assert their right to take pride in and claim responsibility for their collective efforts to reclaim public spaces, engage in protest and direct action, create alternative spaces and epistemologies, and challenge ableist practices and policies. This move-

ment refutes the stereotype of the ‘supercrip’ as someone who, despite their disability, manages to achieve conventional markers of success. Disability pride raises awareness that disabled activists, *because of* their disability, have collectively transformed society in countless ways. What we deserve praise for is so much more than assimilating into capitalist structures. We have rather challenged capitalist systems of ableist oppression, not despite our disabilities but because of them, along with the knowledge and solidarity that they provide.

These examples of disability activism demonstrate that retribution and rehabilitation can be valid and useful forms of responsibility praxis, but only when they serve the higher-order aim of opposing oppression. The ameliorative model keeps this aim in plain sight when we practice responsibility, prompting us to ask whether a given use of responsibility would mitigate or exacerbate oppression in light of entrenched structural injustices. This focus helps us to avoid abstract, decontextualized, ahistorical, and idealistic claims about responsibility that are detached from real-world contexts of oppression and real-life experiences of injustice.

REFERENCES

BELL, Macalester. *Hard feelings: The moral psychology of contempt*. New York: Oxford University Press, 2013.

BELL, Macalester. “The standing to blame: A critique”. In: COATES, D. Justin & TOGNAZZINI, Neal (ed.). *Blame: Its nature and norms*. Oxford: Oxford University Press, 2012. p. 263-281.

CARUSO, Gregg. D.; PEREBOOM, Derk. *Moral Responsibility Reconsidered*. Cambridge: Cambridge University Press 2002.

CLARE, Eli. *Brilliant imperfection: Grappling with cure*. Durham: Duke University Press, 2017.

COATES, D. Justin & TOGNAZZINI, Neal. “Blame”. In: ZALTA, Edward N.; NODELMAN, Uri. *The Stanford Encyclopedia of Philosophy*, 2024. <https://plato.stanford.edu/archives/fall2023/entries/blame/>.

COLLINS, Patricia Hill. *Black feminist thought: Knowledge, consciousness, and the politics of empowerment*. New York: Routledge, 2022.

DAVIS, Angela. Y. *Are prisons obsolete?* Seven Stories Press, 2011.

DORIS, John M. *Talking to our selves: Reflection, ignorance, and agency*. Oxford: Oxford University Press, 2015.

FORMAN, Milos (Director). *One Flew Over the Cuckoo's Nest* [Film]. Warner Bros, 1975.

FRENCH, Peter. A. *The virtues of vengeance*. Lawrence: University Press of Kansas, 2001.

FRIEDAN, Betty. *The feminine mystique*. New York: W.W. Norton & Company, 1963.

FRYE, Marilyn. *Politics of reality: Essays in feminist theory*. Trumansburg: Crossing Press, 1983.

HASLANGER, Sally. "Gender and race: (What) are they? (What) do we want them to be?" *Noûs*, v. 34 , n. 1, 2000, p. 31-55.

HIERONYMI, Pamela. "The force and fairness of blame". *Philosophical perspectives*, v. 18, 2004, p. 115-148.

HOLROYD, Jules. "Oppressive praise". *Feminist Philosophy Quarterly*, v. 7, n. 4, 2021 <https://doi.org/10.5206/fpq/2021.4.13967>.

HOLROYD, Jules. (2024). "The distortions of oppressive praise: Challenges for practice-dependent theories of moral responsibility". *Feminist Philosophy Quarterly*, v. 10, 2024, <https://ojs.lib.uwo.ca/index.php/fpq/article/view/18718>.

HOOKS, Bell. *All about love: New visions*. New York: Harper Collins, 2000.

HOOKS, Bell. *Cultural criticism and transformation*. Interview transcript. Media

Education Foundation, n.d., 1997, <https://www.mediaed.org/transcripts/Bell-Hooks-Transcript.pdf>.

LEVY, Neil. *Hard luck: How luck undermines free will and moral responsibility*. Oxford: Oxford University Press, 2011.

LEWIS, Sophie. *Abolish the family: A manifesto for care and liberation*. London: Verso Books, 2022.

LORDE, Audre. *Sister outsider: Essay and speeches*. Trumansburg, NY: Crossing Press, 1984 https://rhinehartibenglish.weebly.com/uploads/2/2/1/0/22108252/sister_outsider_audrey_lorde_ib_pdf_packet.pdf

LUGONES, María. "Playfulness, "world"-travelling, and loving perception". *Hypatia*, v. 2, n. 2, 1987, p. 3-19.

MANNE, Kate. *Down girl: The logic of misogyny*. New York: Oxford University Press, 2017.

MATTLIN, Ben. *Disability pride: Dispatches from a post-ADA world*. Boston: Beacon Press, 2022.

MARX, Karl. *Das kapital: kritik der politischen oekonomie*. Hamburg: Verlag von Otto Meissner, 1867.

MCGEER, Victoria. "Building a better theory of responsibility". *Philosophical Studies*, v. 172, 2015, p. 2635-2649.

MEISSNER, Shelbi Nahwilet; SMITH, Andrew. Frederick. "Climate crisis as relational crisis: Centering Indigenous feminist conceptions of responsibility in environmental discourse". *Feminist Philosophy Quarterly*, v. 10, 2024, <https://ojs.lib.uwo.ca/index.php/fpq/article/view/18724>.

METZL, Jonathan M. *The protest psychosis: How schizophrenia became a black disease*. Boston: Beacon Press, 2010.

MCKENNA, Michael. *Conversation & responsibility*. New York: Oxford University Press, 2012.

MOORE, Michael. S. *Placing blame: A theory of the criminal law*. Oxford University Press, 2010.

MURTHY, T. Virgil. "Against Arguing about Addict Agency". *Feminist Philosophy*

Quarterly, v. 10, 2024, <https://ojs.lib.uwo.ca/index.php/fpq/article/view/18722>.

NEWNHAN, Nicole, & LeBrecht, James (Directors). *Crip Camp: A Disability Revolution* [Film]. Netflix, 2020.

PATEMAN, Carole. *The sexual contract*. Stanford University Press, 1988.

PITTS, Andrea. J. "Disability bioethics and race". In: REYNOLDS, Joel Michael; WIESELER, Christine (eds). *The Disability Bioethics Reader*. New York, London: Routledge, 2022. p. 35-242.

SMITH, Angela. "Moral blame and moral protest". In: COATES, D. Justin & TOGNAZZINI, Neal (ed.). *Blame: Its nature and norms*. Oxford: Oxford University Press, 2013. doi:10.1093/acprof:oso/9780199860821.003.0002

SOLOMON, Robert C. "In defense of sentimentality". *Philosophy and Literature*, v. 14, n. 2, 1990, p. 304-323.

STRAWSON, Peter. F. *Freedom and resentment and other essays*. London: Routledge, 2008.

TÁÍWÒ, Olúfemi. O. *Elite capture: How the powerful took over identity politics (and everything else)*. Chicago: Haymarket Books, 2022.

TREMAIN, Shelley Lynn. “When moral responsibility theory met my philosophy of disability”. *Feminist Philosophy Quarterly*, v. 10, 2024, <https://ojs.lib.uwo.ca/index.php/fpq/article/view/18720>.

TREMAIN, Shelley. Lynn. “Philosophy of disability, conceptual engineering, and the nursing home-industrial-complex in Canada”. *International Journal of Critical Diversity Studies*, v. 4, n. 1, 2021, p. 10-33.

VARGAS, Manuel. *Building better beings: A theory of moral responsibility*. Oxford University Press, 2013.

WALKER, Margaret Urban. *Moral repair: Reconstructing moral relations after wrongdoing*. Cambridge University Press, 2006.

WALLER, Bruce N. *The stubborn system of moral responsibility*. Cambridge, Mass: MIT Press, 2015.

WOLF, Susan. “Blame, Italian style”. In: WALLACE, R. Jay; KUMAR, Rahul; Freeman, Samuel Richard (ed.). *Reasons and Cognition: Essays on the Philosophy of T.M. Scanlon*. Oxford University Press, 2011, p. 332-347.