# Table of Contents

1. **Nursing Home Abolition: Prisons and the Institutionalization of Older Adult Care**  
   Eva Boodman  
   1–21

2. **Evaluating Biomedical Enhancement: A Non-ideal Approach**  
   Andrey Darovskikh  
   23–33

3. **Content Neutrality: A Defense**  
   Joseph M. Dunne  
   35–50

4. **The Ethics of Knowing the Score: Recommendations for Improving Boxing’s 10-Point Must System**  
   John Scott Gray and Brian Russ  
   51–62

5. **Difficult Circumstances: Situationism and Ability**  
   Marcela Herdova and Stephen Kearns  
   63–91

6. **Nationalism & Social Stability**  
   Stephen McAndrew  
   93–110

7. **Are There Benefits to Benefits?**  
   Holly Stevenson  
   111–123
Nursing Home Abolition: Prisons and the Institutionalization of Older Adult Care

Eva Boodman
William Paterson University

Acknowledgments
I’m grateful to Nathalie Batraville (Dartmouth University) and Tim Johnston (SAGE: Services and Advocacy for GLBT Elders), for their input and suggestions, which helped shape the argument of this paper.

Biography
Eva Boodman is Visiting Assistant Professor of Philosophy at William Paterson University. Her current research focuses on normative approaches to the ways that institutions like schools, prisons, nursing homes, and social service organizations reproduce or mitigate structural racism. Her other projects use feminist ethics, critical race theory and Latinx philosophy to contend with questions of identity, responsibility, and complicity.

Publication Details

Citation
Abstract
This article uses a prison abolitionist model to argue for nursing home abolition. By looking at the development of care institutions during the nineteenth century great confinement of the poor, and the example of black women paroled into forced domestic work at the beginning of the twentieth century, I show that both prisons and nursing homes shape citizenship as participation in a paid, able-bodied, white workforce that finds its condition of possibility in the violent exploitation of people of color in sectors like care work. Because prisons and nursing homes share this historical connection and political rationale, a similar framework can be applied to each: that of abolitionism. The prison abolitionist model shows us that the US needs a transformative approach to older adult care that goes beyond deinstitutionalization. As scholar Ai-Jen Poo writes, creating a society that cares for its elders will require us, as prison abolition does, to “rethink everything – how we live, how we work and play, and especially how we organize our family and community life: how we take care of each other across generations” (Poo 2015, 40). This paper makes the case for that radical re-conception of care by critiquing the carceral legacy of nursing homes.

Keywords
Prisons, Nursing Homes, Abolition, Domestic Work, Race, Citizenship

Introduction
Over the course of the twentieth and twenty-first centuries, mass incarceration in the United States has separated vast numbers of people from their communities and loved ones. Concurrent with this racialized civic death, another vast removal is taking place: the widespread institutionalization of older adults. While there are obvious and important differences between the “prison boom” and the “elder boom”, and between penal and older adult care institutions, both share a role in normalizing civic participation as white and respectable, in opposition to the disposability and exploitability of racialized, gendered, and aging bodies. During the era of U.S. “crimmigration”—the criminalization of migrants, many of whom represent demographics engaged in care work—the relationship between care work and prisons is of particular urgency. By looking specifically at the example of black women paroled into forced domestic work at the beginning of the twentieth century as a precedent for the highly classed and raced care environments we know today, we can see that both prisons and nursing homes normalize citizenship as participation in a paid, able-
bodied, white workforce that finds its condition of possibility in the violent exploitation of black people and people of color in sectors like care work. Because prisons and nursing homes are both sites of this kind of normalization, a similar framework can be applied to each: that of abolitionism. The prison abolitionist model shows us that the U.S. needs a transformative approach to older adult care that goes beyond deinstitutionalization. As care work scholar Ai-Jen Poo has put it, creating a society that cares for its elders will require us, as prison abolition does, to “rethink everything – how we live, how we work and play, and especially how we organize our family and community life: how we take care of each other across generations” (Poo 2015, 40).

The paper begins by describing the current “elder boom” – a term used to describe the rise in the number of older adults – and documenting the problems with the nursing home model of older adult care, both in terms of quality of and access to care, and in terms of labor conditions for care workers. I focus specifically on nursing homes over other kinds of institutional older adult care, both because a fuller, comparative analysis of different institutional models of older adult care would be too much for one paper to take on, and because of those different models, nursing homes have the closest historical and institutional link to prisons in the nineteenth century great confinement. Nursing homes originated in the almshouses of the nineteenth century, which housed older adults alongside those considered mentally ill, vagrant, deviant, or otherwise unable to work. Significantly for the purposes of exploring the connection between nursing homes and prisons, the nineteenth century also saw a movement led by white, church-affiliated women to differentiate older adult care based on “ethnic background”, motivated by the concern that “worthy individuals of their own ethnic or religious background might end their days alongside the most despised of society” (Foundation Aiding the Elderly 2018) – a concern the obvious racism of which reveals the assumption that only those who are respectable and worthy (read: white) are entitled to quality care. The beginning of the paper rejects the widespread use of nursing home care on historical and ethical grounds, motivating the need for alternative models that center community care as part of a widespread agenda for social change that goes beyond deinstitutionalization.

The second section of the paper links nursing home labor to slavery and Southern parole laws that forced black women into domestic labor in what Sarah Haley has called the domestic carceral sphere (Haley 2016, 35). Emphasizing the connection between nursing home labor and violent racial exploitation through penal care work not only makes the relationship between prisons and nursing homes obvious; it also begins to set up the argument that nursing home abolition and prison abolition are not
only related in theory, but should be part of a broader political agenda that reimagines security and care beyond the reliance on institutions that continue to racialize care. The third section suggests that both prisons and nursing homes shore up the boundaries of the state by defining who is entitled to civic participation (versus civic death) according to racialized notions of respectable work. The remainder of the paper moves from these historical and political connections to then use some established tenets of the prison abolitionist approach to sketch a preliminary theory of nursing home abolition. Rather than making specific policy suggestions or outlining what nursing home abolition would look like in practice, I use the analogy between prison and nursing home abolition to motivate further discussion and connection between the movement for decarceration, and the movement for community-based older adult care, and to support the idea that, given the historical connection between prisons and nursing homes, decarceration and nursing home abolition are both necessary for true community elder care.

1. The Elder Boom

By 2050, the number of individuals over the age of sixty-five in need of long term care is projected to grow to 27 million. The network of supportive services like meal programs, affordable homecare, and high-quality senior living communities is unprepared for this unprecedented increase in our aging population. As Baby Boomers age, some are calling this massive increase in the number of people over the age of sixty-five an “elder boom”, which could, if it is not approached with new models, result in a massive crisis.

While nursing home care is widely accepted and might be thought to adequately meet the needs of a growing number of older adults, this is far from the case. Nursing homes are not desirable or reliable care settings. Recent studies show that ninety percent of older adults of all groups want to age at home, or “age in place”, and the reasons for this are clear; nursing home conditions are often deplorable for both residents and workers. In addition to this, the cost of institutional care makes it an untenable option: institutional care comes out to about $84,000 per person, per year. This price tag means that those who are middle-class do not qualify for often sub-par Medicaid-subsidized nursing home stays, but also cannot afford to leave work in order to care for a loved-one full time, or to pay for residential care that will meet their own care needs or those of an aging loved one. Activist-scholar Ai-Jen Poo, founder of the U.S. campaign Domestic Workers United and the author of *The Age of Dignity* (2015),
describes the elder boom as an impending national crisis: in today’s economy, it no longer works to expect family members to leave their jobs to care for aging loved ones full time. Home care seems to be the solution, but while twenty-seven million people will soon be needing care, there are currently only three million home care workers in the entire industry—nowhere near enough to care for all those who need it.

Granted, not all institutional care is made equal, and it may be helpful to make some general distinctions between some of the different forms of long-term care in the US, a context characterized by intense and complex forms of inequality linked to race and class, and neo-liberal ideologies that support private, deregulated care industries alongside an often under-funded public sector. On one end of the care spectrum there are retirement communities which often have independent living sections. These are referred to as private pay, since individuals pay for care out-of-pocket, and in these kinds of facilities support staff are present to help with medication adherence, cooking, cleaning, etc. Older adults who require more assistance with what are called Activities of Daily Living or ADLs, often move into assisted living facilities, and those who require consistent care and monitoring will move into nursing care. Nursing care may also have some number of beds designated for subacute rehabilitation, for example where an older adult would stay while recovering from a surgery. A private pay retirement community will have the fewest regulations and oversight, while a Medicaid (government) -funded nursing facility is subject to more intense state regulation. Nursing homes can be either public or private; for-profit, or non-profit. In the United States, two thirds of nursing homes are investor owned, for-profit institutions (Comondore and Deveraux et al. 2009, 1). While some academic studies tend to show that for-profit nursing homes provide lower-quality care, a large proportion of those studies “showed no significant difference in quality of care by ownership”. In their study on the quality of care in private and public nursing homes, Comondore, Devereaux et al. find that “in the long term care market, in which funding is often provided by the government at fixed rates, both for-profit and not-for-profit facilities face an economic challenge that may affect staffing and other determinants of quality of care” (Comondore and Devereaux et al. 2009, 14). What studies like theirs suggest is that nursing homes themselves operate in a manner and context that is not conducive to providing quality care—regardless of who funds or owns the facilities.

Nursing homes, then, are not working. Poo shows that the current model for older adult care, which assumes nursing home institutionalization, is dysfunctional in terms of providing care to those who need it and compensating those who tend to provide that care—and often with lethal results. In spite of some efforts to improve and reform
conditions, Poo writes, widespread neglect and abuse still occur. According to her, “people confined in institutions, especially those with dementia and other cognitive disabilities, are often physically or chemically restrained, and many are administered antipsychotics”, despite the fatal danger these pose to people with cognitive degeneration. Facilities are often understaffed, especially at night, which makes preventable ailments like malnutrition, dehydration, bedsores, dental and gum disease, and infections common. Conditions for nursing home workers are often deplorable: their work is undervalued, under-supported, and inadequately compensated. The majority of institutionalized older adults die within two years of their arrival (Sidell 1997).

Why do it this way if it doesn’t work? Poo writes that the reliance on institutionalization comes in part from the “fear and discomfort with which we approach aging, illness, disability, and death” in US society, which have led us to put ‘older people somewhere we can’t see them: in institutions’ (Poo 2015, 30). While the “crip” and disability rights movements in the United States have had some successes in the decades-long fight against institutionalization as a discriminatory form of exclusion and marginalization, the movement to create pathways to deinstitutionalization for older adults has not had much political traction until very recently – even though the analyses developed by disability studies apply just as well to the marginalization, warehousing, and social expulsion of older adults. That disability studies analysis which applies so well to the mass institutionalization of elders is heavily informed by the analyses developed by “carceral” studies, which take their cue from French philosopher Michel Foucault to look at the “normalizing” role of institutions like prisons, hospitals, mental health facilities, and reform and residential schools. Warehousing elders, to be sure, doesn’t function in the same way as mass incarceration does. For one thing, mass incarceration and detention in the United States has a primary racializing function that nursing home institutionalization does not; though the histories of penal institutions and crime and punishment in the United States are complex, it is a widely accepted view that post-slavery racialized punishment (like convict-leasing and chain gangs) was the historical condition of possibility for mass incarceration as we know it today.

Nonetheless, racialization and valuations of who was worthy of civic participation were significant in the nineteenth century advent of nursing homes as institutions independent from poorhouses. The very first eldercare institutions assumed those entitled to care to be white and in need of protection and isolation from the other more “degenerate” populations housed in the almshouses. The nineteenth century, then, saw a race and class division in older adult care institutions, where those who
Boodman

shared the “ethnicity” of white women’s groups were “worthy” and could benefit from demographically targeted care, to the exclusion of just about everyone else housed in nineteenth century institutions. While this might seem to support the claim that present day nursing homes are the legacy of an option meant for those with racial privilege (rather than analogous to the prison), I cite the history of older adult care institutions in order to show that divisions in older adult care and caregiving had a role in the racialization of care work and the entitlement to care in a way that has continued into the present.

While I cannot trace the history of older adult care in the U.S. in its entirety here, a couple of milestones are important to note. The nineteenth century church-organized nursing homes described above served only a very small number of people, and in fact, did not last long. By the 1930s, a movement to save money and to abandon “dreaded” almshouses, where conditions were deplorable, succeeded in replacing older adult care both in almshouses and other church-based institutions with small pensions so older adults could support themselves at home – a desirable solution in present-day debates, but impracticable without non-monetary forms of care support. Those who resided in almshouses or other public institutions, however, were not eligible for these pensions, and had to pay for care in private institutions. Significantly, the exclusion of poor and non-white older adults housed in poorhouses from this pension program was the reason for the passing of Social Security legislation. In the 1960s and 1970s the number of both private and public nursing homes grew significantly due to the passage of Medicare and Medicaid, and the increasing stimulation from the private care industry. By 1979, seventy-nine percent of all institutionalized elderly persons resided in commercially run homes (Foundation Aiding the Elderly 2018), many of which provided substandard care. These institutions were “labeled ‘warehouses’ for the old and ‘junkyards’ for the dying by numerous critics” (Ibid.). Representative David Pryor is said to have proclaimed in his attempt to initiate legislative reform in 1970 that nursing homes were ‘halfway houses between society and the cemetery’ (Ibid.).

While this dire picture may not describe each and every institutional care environment for older adults, the advocacy group Foundation Aiding the Elderly (FATE 2018) writes in their history of nursing homes that

the development of the modern-day industry reflects its historical roots. [...] In their initial policies, New Dealers were anxious to sever the connection between old age and pauperism. In barring all residents of public institutions from receiving pensions, however,
they clearly underestimated the proportion of elderly persons who required residential support. [...] The problems that face long-term care for older adults are clearly tied to their historical development. In shutting the almshouse door, policymakers gave birth to the modern nursing-home industry (FATE 2018).

In this sense, nursing homes can be thought of as an outgrowth of the poorhouse, a carceral institution inspired by Bentham’s panopticon that was to serve as the model of the prison and other penal institutions.

Against this historical backdrop, it becomes easier to make sense of recent statistics in nursing home demographics. Since 1999, the number of “minority residents” in nursing homes has increased more rapidly than the minority population overall, and the rise in the number of black nursing home residents has risen by the same amount that the white population has declined (Feng, Fennell, Tyler, Clark and Mor 2011). These numbers indicate – in concert with the historical precedent set in the nineteenth century – that racialized minorities have unequal access to home and community-based alternatives to nursing home institutionalization, making race an intersecting factor in the warehousing of older adults. It’s evident that nursing homes do not currently have the dramatic racializing role that prisons have as a result of the racialization of crime in the United States. Nonetheless, the institutionalization of elders and the institutionalization of criminalized people in the United States do share a common political logic for who deserves care, and a documented historical connection, which, I argue, has important implications for the way we approach contemporary older adult care. In order to understand this shared logic, we’ll have to take an intersectional look at who is being institutionalized, and who is doing the labor of caring for those both inside and outside of institutions.

2. Slavery, Parole, and the Enforcement of Domestic Work

Since the 1980s, the number of incarcerated people in the U.S. has increased from about 50,000 to over two million. Beyond those incarcerated in correctional facilities, one in twenty adults – or eight million people – are under some form of carceral state control, including parole, probation, and other forms of state supervision and surveillance (Gottschalk 2015, 1). The result for most of those who have contact with the criminal punishment system is “civil death”, the denial of basic civil liberties and social benefits including access to housing, education, work, government benefits, political participation, and even access to one’s children (Ibid.). While there are a
number of different explanations for the rise and persistence of mass incarceration in the U.S., I follow scholars who understand it not to have been caused by an objective rise in crime, but by the historical racialization of criminality; the transformed legacy of slavery through the convict leasing system established at the time of the implementation of the Thirteenth Amendment (Kilgore 2017); and the continual repression of people of color movements, through which technologies of surveillance, population control, policing, and other forms of state violence were perfected (Fernandez 2015, xxix-xxxvi). Political scientist Marie Gottschalk has challenged this explanation, writing that it is a “mistake to view the carceral state as merely the latest chapter in a book that began with slavery”, and that we must, conversely, look at how periodic moralistic political campaigns, and institutions of social welfare (or the lack thereof), shaped criminal punishment. While Gottschalk is right that the history of penal institutions in the U.S. is complex, and that prisons cannot be equated in any simplistic way to slavery or Jim Crow, the role of penal institutions of the past in creating the racializing system that we know today cannot be denied. This section looks at nineteenth century parole laws that forced incarcerated black women into unpaid care labor in white homes. This historical example of racialized criminal punishment represents a direct line of connection to the contemporary demographics of care work: most of those employed as care workers, both inside and outside of institutions, are women of color; many are immigrants or undocumented.¹

In a society afraid of death, illness, and disability, those who must perform labor related to it are devalued, feminized, considered disposable, deportable, and exploitable. Paralleling the exclusion of poor and non-white older adults from care institutions, as we saw above, care workers have been excluded from labor laws in the US since the New Deal – an exclusion that isn’t simply a similarity that domestic and care workers have with incarcerated workers, but that is historically rooted in the legacy of slavery. As Poo writes, “farmwork and domestic work had been the work of slaves, so when labor laws were passed in the 1930s, southern members of Congress refused to sign on to the labor law package as part of the New Deal if farmworkers and domestic workers—who at the time were largely African American—were included. In the deal that Congress struck, those two workforces were excluded” (Poo 2015, 88). It wasn’t until 2013 that “care companions” were included in the Fair Labor Standards

¹. U.S. Census data shows that 89% of home care workers, for example, are women, and that while women of color make up one forth of the workforce, they make up more than half of the home care workforce (“U.S. Home Care Workers: Key Facts”).
Act, covering two million care workers under minimum wage and overtime protections that were implemented in 2015, with extremely uneven enforcement; care workers continue to face the legacy of slavery in the *de facto* devaluation and stigmatization of their work, even though there have been some *de jure* improvements.

Sarah Haley’s scholarship on the nineteenth century use of parole to force black women into care labor in the U.S. South goes further than simply explaining why domestic workers are denied labor rights; it reveals the history of care work in the U.S. as itself a carceral environment, where care work was racialized because of its role as a penal technique. Far from merely punishing “crime” or “criminality”, Haley writes that this form of gendered labor exploitation was a form of neo-slavery that “made the New South possible, not as a departure from the Old, but as a reworking and extension of previous structures of captivity and abjection through gendered capitalism” (Haley 2016, 2)—a system that served to maintain white women’s social and economic role as domestic “managers” of black women in order to support and normalize a white male workforce (Haley 2016, 4).

In 1908, the State of Georgia replaced convict leasing with the chain gang, and, shortly afterwards, the law was amended to “exclude females” from chain gang work. But, alongside this change, parole law was modified so that the Georgia Prison Commission was now empowered to conscript women prisoners to work for private individuals or businesses “for at least one year” instead of releasing them after their minimum sentence (Haley 2016, 157). When this year of unpaid domestic servitude supervised by private white individuals had been completed, there was no guarantee of pardon or commutation (Haley 2016, 175). The result was the “dramatic entrenchment” of black women in domestic work: by 1890, four out of five black women in Georgia were domestic workers; by 1930, this was true for nine out of ten black women (Haley 2016, 34).

While these new laws did affect some white women, the 1913 Probation Act saw the continuation of the differential treatment white and black women received in the penal system. These new probation laws authorized judges to reduce felonies to misdemeanors, and did not require those being released from prison or penitentiary

---

2. Haley writes that ‘domestic service parole’ had an important role in defining gender through race. ‘Georgia’s Jim Crow carceral regime’, she writes, ‘produced women every day, and all of the women were white’ (Haley 2016, 160). Black women – who were both sentenced to work on the chain gang, in the fields, and in white homes – were gendered in ways that were ambiguous, and that acted as a condition for innocent, vulnerable, white Southern femininity.
to live with their employers so long as they reported to a probation officer and led a “correct life”. Probation was mostly reserved for white women, who were more likely to be charged with misdemeanors rather than felonies, while black women were disproportionately given domestic labor parole, which required women to live with their employers, mostly private individuals who had volunteered to act as parole officers (Haley 2016, 34). Black women paroled into unpaid domestic service for white employers were “always under threat of being sent back to the chain gang or state farm if they broke a rule or failed to work up to their employers’ standard” (Haley 2016, 176). Employers had to endorse parolees’ monthly reports, and were meant to “direct her in that which is good, and promptly report to the Prison Board any unnecessary absence and tendency to low and evil associations” (Ibid.). In order to have access to this free domestic labor, white men and women competed for incarcerated women, asking for them to be released into their household’s custody, and often invoking – especially in the case of middle-class or upper-class elderly white people – “frailty”, vulnerability, and the need for assistance and care (Haley 2016, 181). This is not to deny that many – as we saw above – are in need of home elder care today, or that many find that model to be preferable to institutionalized care. But because nursing homes as we know them today developed shortly after the abolishment of convict leasing, we can say that carceral models of domestic care had an impact on shaping a care labor economy based on racial exploitation.

Haley’s work shows that the punitive nature and racialization of domestic work were the condition for the advancement of white men and women at the expense of black women, whose supposed criminality and moral degeneracy allowed for, and was juxtaposed against, white feminine frailty and innocence. While this picture of private, state-sanctioned domestic carceral punishment might seem to support an argument against homecare in favor of institutionalized elder care, the history of care for members of the white propertied class by unpaid, criminalized black women, shows, rather, that prison and nursing home care are inextricably linked. Looking at the way that white middle-class men and women in the South asserted their “right to ‘have’ black women’s domestic labor and bodies” and rationalized their requests as “liberal efforts to liberate them from prison and to care for them” makes it clear that elder care without a transformative movement to decarcerate will only allow for the continued expansion of the carceral state through our care institutions. We need
to look beyond de-institutionalization to broader transformative models that are not rooted in the racializing, punitive histories of both the nursing home and the prison.³

3. The Normalization of White, “Respectable” Citizenship

The practice of domestic labor parole is an example of the racialization of care work as a direct outgrowth of slavery, convict leasing, and other forms of hyperexploitation that served the interests of the white propertied class and created the conditions for early twentieth century industrialization. While it is true that the lived experiences of care workers in general cannot be directly compared to those of people who have survived incarceration, and that racialized punishment is a central function of prisons in a way that it is not in nursing homes or for nursing home workers, both institutions have a ‘normalizing’ function that at once rationalizes the exploitation of people of color and shores up the notion of white citizenship engaged in “respectable” work.

“Normalization” is a term drawn from the work of Michel Foucault that describes disciplinary techniques that work on the soul rather than the body – a development that marked the transition to modernity and capitalism from feudal times, where the authority of the king was enforced through public execution. In *Discipline and Punish*, Foucault describes the role of prisons *not* as a means of hyperexploitation and civil death, which seems to put him at odds with scholars like Haley, but as a means of ‘rehabilitation’ for those who do not fit the norm of production and obedience. Confined bodies are controlled, not with a view to elimination or exploitation *per se*, but as a means of social control, where punishment is used to correct social deviance. As Foucault writes in *Discipline and Punish*, penal labor, “with its limited extent and its low output, cannot have a general effect on the economy. It is intrinsically useful, not as an activity of production, but by virtue of the effect it has on the human mechanism”—the effect of transforming the convict “into a part that plays its role with perfect regularity” (Foucault 1995, 242).

³ Marie Gottschalk argues in *The Prison and the Gallows* that prisons and care institutions have an inverse relationship to one another in the United States. She maintains that victims’ rights groups would not have had so much power to politicize punishment if the U.S. had a stronger network of care institutions that construed crime victimhood as an issue of care rather than an issue of criminal punishment. While she does accept that crime and punishment had a significant role in defining US nationhood, she claims that the “institutional environment” shaped criminal punishment, rather than conversely understanding care institutions to have been shaped by the carceral state. This paper takes a slightly different approach by understanding institutional life in the United States to be radically shaped by racialized punishment and the legacy of racial exploitation.
Foucault’s view is generally criticized for neglecting the role of race and racial hyperexploitation in the U.S. carceral state. I do think, though, that the notion of normalization helps us understand the connection between prisons and nursing homes, and supports the abolition argument. Where prisons and care work in the U.S. have their origins in patterns of racial exploitation, the warehousing of older adults in nursing homes seems to serve the role of elimination and disposal rather than enforcing productivity or “rehabilitating” deviance. But beyond the differences in their respective functions, prisons and nursing homes both have a normalizing role in defining civic participation as raced and classed that can be traced back to the 19th century’s nation-building efforts. By looking at prisons and nursing homes together, we can see how Foucault’s notion of modern capitalist discipline “coincided in post-emancipation America” with brutal racial terror and exploitation, serving to “normalize white subjects rather than to rehabilitate black ones” (Haley 2016, 27). In this sense, prisons and care institutions like nursing homes have worked in concert to create a racial caste system in the U.S., where racialized subjects are simultaneously disposable and exploitable. That caste system racializes exploitative care work and whitens access to quality care as part of the maintenance of a racial-capitalist nation state.

The logic of disposability in the normalization of white citizenship is easy to trace in the twentieth century history of mass incarceration, which some scholars have understood as a tool for punishing those who cannot comply with the demands of the protestant work ethic, as was the case for black workers left unemployed and deemed economically disposable after mass mid-century de-industrialization in the U.S. (Fernandez 2015, xxi-xxxvi). Mass incarceration was a way to warehouse redundant labor, but most of all, it was a way of containing and pre-empting the potential uprisings and “unruliness” of those without work. In this sense, prison wasn’t only a mechanism of exploitation, but was a means of social control that “became necessary after the crisis of the urban ghetto (provoked by the massive loss of jobs and resources attending deindustrialization) and the looming threat of Black radical movements” in response to widespread poverty and deprivation (Ibid).

It is certainly not difficult to see this logic – the logic that punishes redundancy – at work in the institutionalization and neglect of older adults, and not just in the way that U.S. labor culture punishes people for leaving work to care for their loved ones: While we are living longer than ever before in the US, we are also one of the most ageist societies on the globe. The United States is far behind countries like Japan,

---
4. In a 2006 study of perceptions of beauty, for example, the soap company Dove asked women worldwide...
Germany, and Canada, which have created long-term social insurance programs that cover in-home care and community-based solutions that do not penalize those who are poor or without family\(^5\) -- or, at least not to the same dramatic extent. The reason for this, according to sociologist Chris Chapman, is that institutions like nursing homes share a logic that originated with a great confinement of the poor in the eighteenth and nineteenth centuries meant to shore up the boundaries of young nation states by eliminating those who disrupted the vulnerable social order, especially through non-work and “vagrancy” (Chapman 2014, 33). By the nineteenth century, a wide range of different kinds of groups were institutionalized on this logic, which, with the process of secularization, saw a shift from the goal of elimination to the goal of “reformation” or “rehabilitation”. These different institutions shared the political rationality that “under the right conditions imposed from above, degenerate, disabled, criminalistic, or uncivilized peoples” could be “brought ‘up’ to normative standards”—that is, the ability to work (Ibid). “Slaves, First Nations, paupers, criminals, or intellectually, physically or psychiatrically disabled people” were institutionalized in various settings meant to integrate them into society as menial laborers by eliminating their non-productive deviance. In colonial “secular” nation states like the U.S. where direct elimination of non-Christians and “vagrants” could no longer be justified, elimination could now be achieved through “transformation”, as exemplified in the nineteenth century U.S. Indian Commissioner William Jones’ goal to “exterminate the Indian but develop a man” (Chapman, Carey and Ben Moshe 2014, 6). When reading accounts like Haley’s, where a similar rationale was used by white middle-class people vying for the unpaid labor of paroled black women, we can see that “elimination through transformation” was synonymous with racial exploitation under the guise of “rehabilitation” and white liberal paternalism.

This logic of “elimination through transformation” might not seem applicable to elder care, since elders are presumed not to be able to be “rehabilitated” and exploited for manual labor. But nursing homes participate in this political logic since their elimination of older adults is undertaken for the sake of preserving a social order

\(^5\) These programs took decades to implement, with significant adjustments and improvements over time.

‘in which decade of life they thought a woman reaches the peak of her beauty’. Forty-eight percent of American women said ‘in their twenties, while the majority of French, Italian, and Brazilian women said a woman is her most beautiful in her forties’. Other reports reveal that eighty percent of older Americans have encountered ageist stereotypes, including employment and health care discrimination on the basis of age. (Poo 2015, 29-30)
governed by norms of productivity and perfectibility, where independence and dignity are thought to be earned through forms of “respectable” work that exclude care work. Chapman writes that it is self-evident that “non-productive” older adults and those who are incarcerated should be segregated away from the rest of society on the prevailing normative political rationality where to be a civic participant assumes being white, respectably productive (involved in paid, non-care, non-domestic work), and able-bodied. “Every time a group home is built”, he writes, “every time someone can’t imagine a world without prisons or psych wards, every time funding is available for a nursing home but not for care in one’s own home, the fundamentals of the political rationalization of the ‘great confinement of the poor’ lives on” (Chapman et al. 2014, 6).

In addition to such a political rationality living on in both nursing homes and prisons, those institutions themselves – through their techniques and mechanisms – generate our ideas of who is a productive contributor, and who can be done without. Likewise for the institutionalization of disabled people, which was so formative in the development of nursing homes: as Abbas and Voronka write, it was within spaces of confinement that “mad people were monitored, typified, and categorized”, and that “‘mental illness’ and ‘the mentally ill’ were able to emerge [as categories] through examination” (Abbas and Voronka 2014, 122). In the same way, nursing homes and institutional environments for older adults have a role in normalizing the idea that our elders are a burden, are objects of medical technology, rather than being contributors to our social, cultural, emotional and civic lives. In much the same way that the language of a “prison boom” normalizes the flow of bodies into a carceral system, language of the “elder boom” or “silver tsunami” plays into fears that older adults are a wave about to crash into our systems, burdening our economy and sucking out our wealth. Studying the logics at work in prisons, nursing homes, and other institutions that operate on the underlying notion that viable work or recognized productivity grant freedom and civic belonging, we can see that carceral logics are fundamentally political and bio-political: they serve to control who is an acceptable or desirable member and participant in the state through the physical removal of bodies not doing the right kind of work – bodies that include those of undocumented people, care workers, incarcerated people, and older adults.
4. Beyond Deinstitutionalization: Nursing Home Abolition

An obvious objection to the idea of nursing home abolition is that eliminating the institution would place many elders at great risk, especially those who rely on nursing homes for the intensive or memory care they need in order to continue living. While we are facing an impending crisis because of the shortage of adequate care, some might say that eliminating nursing homes would create another, much graver crisis for those currently institutionalized, and their family members who cannot afford to care for them full time. But nursing home abolition is not about eliminating the physical site of the nursing home, in the same way that prison abolition is not about simply eliminating the physical site of the prison. Many have levelled similar critiques against prison abolitionists, arguing that prison abolition understood simply as deinstitutionalization will create a safety crisis if there is nowhere to “put criminals”.

But if, as we saw, prisons and nursing homes are sites of normalization that reinforce the notion that civic and social participation require “respectable” productivity, solving the problem will require much more than deinstitutionalization – even if that should be part of the picture. Foucault writes that the techniques that operate within the prison developed as part of a massive societal shift where institutions, including nursing homes and prisons, played the role of enforcing productivity, docility, and non-deviance through the principles of isolation, work, and the modulation of punitive techniques (Foucault 1995, 231–246). What this means is that to shift their normalizing power, we will have to target not just the institutions of prisons and nursing homes, but the sets of norms, values, and auxiliary structures that support them and that are continuous with them. Nursing homes, for example, are part of a constellation that includes ageist attitudes, work expectations and demands, the cost of living, the feminization and marginalization of care workers, and the medicalization of old age. Likewise, prisons are part of a constellation that includes the prisons themselves, police and police practices, the construction of race and racialization through the history of labor and the legacy of slavery, access to work and education, and the operation of white supremacy at work on a number of different registers of social and political life. Such a wide-ranging and systemic problem demands an equally wide-ranging and systemic approach, though there are evidently no easy solutions to the structural racism, classism and ableism perpetuated by both prisons and nursing homes. The abolitionist approach to both prisons and nursing homes responds to the fact that deinstitutionalization on its own –while an important part of the strategy – will not shift the constellation of norms, practices, and historical brutalities that give prisons and nursing homes their role.
Prison abolition has been defined as “a political vision with the goal of eliminating imprisonment, policing, and surveillance and creating lasting alternatives to punishment and imprisonment”—on the grounds that prisons are a racialized, and inhumane form of punishment and social control (“What is the Prison Industrial Complex?”, criticalresistance.org). In their mission, the leading U.S. prison abolitionist organization Critical Resistance writes that abolition isn’t just about “getting rid of buildings full of cages”. It’s also about shifting the norms, patterns and structures at work the society we live in “because the prison industrial complex both feeds on and maintains oppression and inequalities through punishment, violence”, and social control. In their description of an abolitionist vision, Critical Resistance takes abolition to require a broad set of different strategies, since the prison industrial complex is “not an isolated system”, but a vast network of punitive mechanisms of state violence that includes the institutionalization and social exclusion of older adults, people of color, people with disabilities, queer people, and others who do not conform to the image of the productive citizen.

While there are serious debates about how to carry out abolitionist work, prison abolitionists tend to agree with scholar-activist Mariame Kaba that it is both “lodestar and a practical necessity” (Kaba, Berger, and Stein 2017; both a set of practical organizing tools, and a set of long-term goals for social and political transformation. As Kaba has written, abolition is about “connecting a radical critique of prisons and other forms of state violence with a broader transformative vision” (Ibid.). But is abolition just a “utopian fever dream”, as anthropologist Roger Lancaster has claimed (Lancaster 2017)? Against Lancaster’s skepticism, Kaba and others have argued that prison abolitionism is a highly influential approach that has resulted in significant changes for those affected by mass incarceration. An advantage of abolitionism, according to Kaba, is that “rather than juxtapose the fight for better conditions against the demand for eradicating institutions of state violence”, it navigates that very divide—and that navigation has had some important movement results (Kaba et al. 2017). For the better part of the last fifty years, Kaba writes,

abolitionists have led and participated in campaigns that have fought to reduce state violence and maximize people’s collective wellbeing. They have worked to end solitary confinement and the death penalty, stop the construction of new prisons, eradicate cash bail, organized to free people from prison, opposed the expansion of punishment through hate crime laws and surveillance, pushed for
universal health care, and developed alternative modes of conflict resolution that do not rely on the criminal punishment system (Ibid.).

As Ruth Gilmore has written, prison abolitionism is not meant to be an all-or-nothing principle, but rather is an approach that rejects the paradigm where prisons serve as catchall solutions to social problems (Gilmore 2007). Some conceive of abolitionism as opposed, or antithetical, to reform efforts, but for most abolitionist advocates, this is far from the case. Rather than being opposed to incremental change, “abolitionists have insisted on reforms that reduce rather than strengthen the scale and scope of policing, imprisonment, and surveillance” (Kaba et al. 2017).

Prison abolition advocates, while they focus on prison, have mobilized not only against prisons, but against policing, the school to prison pipeline, cash bail, solitary confinement, stop-and-frisk, punitive benefits requirements, and the current functioning of the foster care system. To affect the way racialized punishment operates in society – or the “punishment paradigm”, as some have called it – requires a widespread cultural, material, relational, policy, and labor shift (Farid 2015, 5–9). It would require making changes to housing laws, zoning, and how school is funded, and rethinking who does what kind of work in society. Alongside more “incremental” efforts, abolitionism also requires a thoroughgoing critique of the nation state and the operation of government – in keeping with the above analysis of the great confinement of the poor as a way of enforcing respectable work as a condition of citizenship and personhood.

The same paradigm shift advocated by prison abolitionists can easily be applied to nursing homes. While there are some important differences in the way these institutions function – for example, jobs in “corrections” are socially sanctioned, and are protected by powerful unions with a good deal of political clout, which is not the case for domestic workers – the argument that abolition involves a transformative shift beyond deinstitutionalization holds for nursing homes, too.

Just as prison abolition does not advocate only for the elimination of prisons, deinstitutionalization on its own will not solve the problem of the elder boom. As Poo’s work shows, we need a shift in policy, relationships, and culture – changes that are unlikely to happen in the short term, and will require incremental organizing efforts. Scholars in disability studies have made a similar argument about deinstitutionalization. While institutions, have had, as we saw, an important role in the construction of “the criminal”, “the mad person”, and “disability” that can be applied also to “old age”, deinstitutionalization would not be the ultimate or singular goal of the
abolition of nursing homes. Disability studies scholar Michael Rembis has shown that deinstitutionalization for those considered mentally ill has, in fact, had the effect of increasing mass incarceration, since prisons have come to replace the ‘correctional’ role that mental institutions once had (Rembis 2014). Deinstitutionalization in the case of elder care – without other concurrent changes – would likely have similar results: an increased number of elder homeless who end up incarcerated or confined in police detention, jail, or in costly hospital stays not intended for long-term care. Because of this, nursing home abolition, on analogy to prison abolition, would mean a long-term, policy and community-supported shift in how we live and work. In the same way that prison abolitionists seek to redefine our notion of safe communities, nursing home abolition would involve a creative redefinition of community care that goes beyond nuclear families and state institutions as the only sources of support.

For communities of color, immigrant communities, and queer communities, informal community-based care structures have, at some times, been a cultural norm, by necessity or by choice. In their paper “African American extended families and kinship care”, for example, Stephanie Brown, Don Cohon, and Rachel Wheeler suggest that care structures based on extended family networks that include non-biological family are often “disrupted” by state-managed care services, and that “continued idealization of the nuclear family—including its use in the conceptualization of foster care—may hinder service provision because it obscures the resources of extended families” (Brown, Cohon, and Wheeler 2002, 53–77). Non-institutional, non-nuclear family-based structures of care have historically been disrupted by state violence and repression that break extended families apart through institutionalization and the enforcement of colonial models of family. Nursing home abolition is a response to the colonial, carceral models of care that enforce racialization, and that, in this stage of capitalism in the U.S., are not working. Nursing home abolition, on analogy to prison abolition, would mean a long-term shift in how we live and work, in who we expect

6. While queer elders in particular face barriers when accessing care that include the assumption of heterosexual identity, lack of recognition of same-sex partners and unequal treatment by providers – in addition to a history of medicalizing and pathologizing queerness -- the way society cares for elders (and people convicted of ‘crimes’) has a lot to learn from queer models of care that decenter biological family, lean on community, are multi-generational, and mistrust care institutions. A group or community setting can be the best environment to provide safe and necessary care for an older adult. For example, providing memory care is extremely stressful and demanding. The repetitive behaviors, wandering, confusion, and aggression of older adults with dementia can be very hard to handle, especially for someone not trained in dementia care, putting the elder at greater risk of neglect or elder abuse.
to care for, and who we expect to care for us. Understanding the deeper historical and political connections between prisons and nursing homes shows us that transformative, incremental changes in criminal punishment and elder care must go hand in hand. Without a movement for decarceration, we will not have true community care for our elders.

References


“What is the Prison Industrial Complex? What is Abolition?” In criticalresistance.org.


Feng, Zhanliang, Mary L. Fennell, Denise A. Tyler, Melissa Clark, and Vincent Mor. 2011. “Growth of Racial and Ethnic Minorities in US Nursing Homes Driven by


“The History of Nursing Homes.” In *Foundation Aiding the Elderly*.


“Home Care Workers: Key Facts.” In www.PHInational.org.


Kilgore, James. August 5th, 2017. “Will Altering the 13th Amendment Bring Liberation to the Incarcerated 2.7 million?” In *Truthout.org*


Evaluating Biomedical Enhancement: A Non-ideal Approach

Andrey Darovskikh
Binghamton University

Acknowledgments
An earlier version of this paper was presented at the Center for Cognition and Neuroethics’ Ethics and Social, Political, and Legal Conference (Flint, MI) in November 2018. I am very thankful for the provided feedback.

Biography
Andrey Darovskikh is a PhD Candidate at Binghamton University. He holds MAs in Philosophy and Interdisciplinary Medieval Studies, and specializes in ancient philosophy with a particular focus on ancient biology and medicine. He has a developing interest in applied ethics, especially in bio and medical ethics.

Publication Details

Citation
Abstract
The problem of how to assess technological capabilities of enhancing human nature and health has been haunting for scientists, philosophers and general audience for a significant period. However, in this paper, the author suggests to leave aside the question of normativity with regard to biomedical interventions. Instead of that, the focus is the evaluation of a particular act of enhancement provided that human enhancement in itself is a morally right thing to do. The discussion focuses on the debate about the retrospective (backward-looking) and the prospective (forward-looking) evaluations. On the one hand, the author verifies the validity of the argument that the backward-looking approach suffers from shortsightedness. On the other hand, the author objects to claim, that the only alternative to the backward-looking approach is to look forward to achieving an ideal, as if it is indispensable for qualifying both quantitative and qualitative chances as positive improvements. As a response, the author proposes an approach for evaluating human enhancement, which is both non-ideal and forward-looking at the same time. In the center of the argument is the idea that we can adequately evaluate a particular act of enhancement by comparing the improved state with some descriptively and normatively better state, which is not an ideal one. This better state is non-ideal, because it is a subject to revision depending on human nature, social circumstances and availability of technologies.

Introduction
The goal of human enhancement is the pursuit of a better life. In the course of the history, people always tried to improve themselves: illiteracy eradication; urbanization; vaccination; development of technologies (e.g., we are flying, computing and the like). However, biomedical enhancement is a different case because of its potential power to affect efficiently human bodies and brains in so many ways and often times so quickly. In order to qualify a change as enhancement we need to look at the evaluation of its influence on the subject and answer the question whether we perceive the change as increasing functionality and capabilities or not. Biomedical enhancement is no exception in this regard. Within the scope of this paper, I will leave aside the problem of normative evaluations of biomedical enhancement in general. The ethical pro & contra debates about biomedical enhancement have been going on since science provided (or promises to provide) us with a means to improve human nature.
more drastically than ever. However, let us for the sake of the argument allow the premise that biomedical enhancement in general is a morally right thing to do; not in the sense that it is something obligatory but it is at least permissible and at most advisable. Then the question arises; how do we evaluate every particular act of human enhancement? There are two ways of evaluating improvements in non-normative terms: 1) Retrospective evaluation (frequently referred to as backward-looking), which suggests assessing how far we managed to advance from some original position. 2) Prospective evaluation, which assesses how close we approached the state, which was considered as a goal of the improvement.

The retrospective approach assesses enhancement looking back to status quo ante. The status quo ante is not some kind of a deficient state, nor is it an average condition, but it refers only to the way things used to be. In biomedical enhancement, it would be a physical or mental condition of a particular individual prior to biomedical intervention. Given the knowledge about that particular unfavorable state from the past, enhancement can be assessed based on how successfully those unfavorable conditions were overcome. On the other hand, there are several ways to talk about forward-looking evaluations, and one of them suggests a necessity of some kind of ideal ahead. Johann Roduit one of the proponents of this approach argues that an ideal seems indispensable when a comparison of an original and an improved state is needed. For this type of a forward-looking approach an ideal as a long-term goal is necessary because it allows qualifying either quantitative or qualitative change as a real improvement in case if it contributes to attaining the ideal (Roduit 2016, 71).

In this paper, I express some skepticism about the ideal forward-looking approach. I object to the claim that current circumstances with regard to human nature cannot influence the direction of enhancement. I will elaborate this position by expanding the scope of current circumstances to include not only the state of human nature but also social and technological factors. This does not imply that I advocate the backward-looking approach. On the contrary, I will verify the validity of the argument that this approach suffers from shortsightedness. However, I am going to refute the claim that biomedical enhancement cannot be properly understood without a reference to an ideal. I will propose an approach for evaluating human enhancement, which is both non-ideal and forward-looking at the same time. On the one hand, I will stick to the common argument from non-idealists that an ideal is not necessary for deciding which of two states (original or improved) is more favorable. On the other hand, I affirm that what matters here is not a comparison of an improved state and a status quo ante, but a comparison of improvement with some descriptively and normatively better state,
which is not a perfect or an ideal state. Finally, I will show that this better state is non-ideal, because it is open to revisions and depends on current individual and social circumstances and especially on availability of technologies.

The Case of Biomedical Enhancement

Human enhancement always aims to ameliorate unfavorable conditions. The question what is an unfavorable condition in human nature or health is problematic due to plurality of views and some ethical constrains. On the one hand, individual or cultural subjectivity suggests that the possible list of unfavorable human conditions has no limit. For example, small breasts or insufficiently long erection are unfavorable conditions for many by far; bad hearing or sight are as well; low IQ or inability to sleep three hours per night while being efficient during the day – all those states are possible targets of biomedical enhancement. On the other hand, different moral constrains force us to be careful with labeling that or another physical state as unfavorable on a large scale. For example, the arguments in favor of preventing the eradication of deaf culture are well known (See: Chadwick and Levitt, 1998).

The amelioration of an “unfavorable” condition, which restores the state of health and normal species functioning, is frequently referred to either as negative enhancement or therapy. Positive enhancement broadly speaking is a desire to get from good to better. It is a pursuit of improving the normal human functioning for the sake of getting an upper hand. In order to give a full account of the problem of biomedical enhancement the distinction between negative and positive enhancement should be multiplied by the distinction between somatic and genomic (or germline) interventions. If we combine this distinction, the full list would include four types of biomedical enhancement: 1) Somatic negative enhancement 2) Germline negative enhancement 3) Somatic positive enhancement 4) Germline positive enhancement.

The understanding of the distinction between somatic and germline interventions hinges on the understanding of the difference between somatic deviations (mutations) and germline deviations (mutations). Somatic deviations occur in a single body cell, which means that they affect only this particular individual and they cannot be inherited, because only the substance between derived deviated cells is affected. On the contrary, in the germline mutations the deviations occur in the gamete and will be transmitted to the offspring of this particular individual. The characteristics acquired in the process of this intervention will be inherited because the entire organism will be affected.
Despite the drastic relevance of the distinctions between different types of biomedical enhancement for debating the normative status of biomedical interventions, they seem to be less significant for the purposes of my argument, because the criteria to be used for evaluating particular enhancement are the same the type of biomedical intervention notwithstanding. Nevertheless, I do want to set a couple of limits for the kinds of biomedical enhancements, which will be in the focus. First, biomedical enhancement is something available only due to the development of biomedical science, i.e., limbs amputation is not biomedical enhancement, even though some might argue that life without legs is a desirable and more beneficial state. Second, enhancement can be instrumental or even indulging pleasures, as it in the case of prolonged erection, but enhancement can be practical and at the same time can have some intrinsic value. I mean that a certain enhanced state is conventionally desirable by a significant fraction of population as it is, in a particular set of historical, social and technological circumstances. For example, the constitution of our backbone is such that excessive sitting literally kills us, although it was not a problem even a hundred years ago when the majority of human kind population would move more. Thus, an enhancement of the spine’s constitution in the nowadays conditions would be a good example of such a conventionally desirable state which is on the one hand instrumental; on the other hand it is intrinsically valuable. These desirable conditions vary immensely depending on time location and specific population but in each case, they target something that is generally perceived as an increase in well-being.

The Critique of the Backward-Looking Approach

The backward-looking approach to evaluating biomedical enhancement has two major problems: 1) it positively evaluates any improvement; 2) it does not take into account a distant perspective of improvements. Evaluation of enhancement in reference to an original state is problematic, largely because a quantitative improvement does not necessarily mean a qualitative improvement. If we consider enhancement as something more than just a restorative therapy, the backward-looking evaluation will be positive every time we depart from the original state and somehow improve a human being. In line with this logic, any enhancement is supposed to always make us better off. The problem of such a view can be easily displayed. Let us imagine that deafness is not desirable by anyone and we all opt for its eradication. Then, enhanced hearing of someone who was originally hearing-impaired is definitely an improvement of a life quality. However, drastically enhanced hearing of the same person makes him no
longer better off. Clearly, the ability to hear every single flap of the butterfly’s wings makes life unbearable.

Furthermore, even a moderate enhancement of some functions, which seems to be beneficial at first sight, might turn out to be detrimental in a long-term perspective. First, some enhanced traits might not be desirable in the near future. If we look back, we can see that fencing skill or horse riding are barely desirable by majority in the 21st century, however it used to be so for a long time in the past. By the same token, something we long for now can become worthless in the future. Second, some types of enhancement can impede the enhancement of more important faculties. Here we can think about side effects of enhancement that interfere with both current faculties and faculties to be enhanced in the future.

Alternatives to the backward-looking evaluating approach suggest that its shortcoming primarily stems from the lack of an envisioned state lying ahead. This desirable state is used as a point of reference and is supposed to steer the whole process of enhancement toward a certain goal. In this paper, I oppose the view that the only way for biomedical evaluation to be forward looking is to refer to an ideal/perfect state. For this view, any kind of an unfavorable condition is a state of imperfection, which is simply caused by a lack of perfection. According to this view as long as we get rid of all imperfections, we achieve the state of ideal perfection; seemingly, without this ideal state, our desire to improve ourselves by means of biomedical enhancement cannot be properly understood.

In the interest of fairness, I need to note, that the more recent defenders of the necessity of perfection as an ideal state such as Johann Roduit put forward a recalibrated and more argumentative update of perfectionism theories in connection to human enhancement feared by, for example, Michael Sandel and Leon Kass (See: Sandel 2007; Kass 2003). According to that view debunked by the latter, the ultimate goal of enhancement is to acquire total mastery, perfection and even immortality. However, our desire to live better and to be better off does not imply a desire to achieve total mastery. Allen Buchanan was one of those who argued that such perfectionist claims have no evidence for support (See also: Caplan 2009, 201) and they are simply wrongheaded, because “the pursuit of biomedical enhancement is not the pursuit of perfection; it is the pursuit of improvement” (Buchanan 2011, 2).

While the title of Roduit’s book (The Case of Perfection) explicitly refers to the idea of perfection and though he always uses the word perfection to describe the ideal state, doubtless his understanding of perfection is in fact not identical with total mastery and immortality. Rather, his perfect human being is an ideal, which complies
with a broad range of conditions necessary for a good life. As he claims, “It makes perfect sense to argue that having or being enough can well be a viable factor for determining human perfection” (Roduit 2016, 82). Roduit is reticent on the point how an ideal human being would look like, he simply does not see his task in talking about a precise ideal. He aims to solve a methodological issue and prove that we do need an ideal when we venture upon biomedical enhancement.

It might even seem that Roduit argues in favor of something other than an ideal. However, I insist that what he defines as a threshold for human well-being (Roduit 2016, 82) constitutes an ideal theory for an ideal human. It is this ‘imperfect ideal’ which is supposed to be taken into account for directing the pursuit of enhancement in any given circumstance, which is non-ideal. For Roduit, there has to be a suitable perfection as an ideal, and people committed to that ideal use biomedical enhancement as a tool to achieve it regardless of any present set of conditions. I argue that what makes Roduit’s theory an ideal one is the lack of interest to the present biological state of humans, the state of the society, and especially the development of science.

**The Forward-Looking Non-Ideal Evaluation**

Let us imagine that the minimum threshold of an ideal life among other features includes a highly developed intellectual capacity. Given that mistakes, misfortunes and sometimes crimes reflect the ability to do something with deep convictions based on no information at all or false knowledge, the idea of increasing IQ seems to be a very desirable ‘enhancement package’. Let us again imagine that biomedical science does have right now this kind of enhancement available. If we perceive human enhancement as something that aims to attain an ideal then this particular enhancement of intellect should be positively evaluated, because it brings us closer to the ideal. However, the highly developed state of intellectual capacity increases the risk of dissatisfaction, which stems from inability to find a suitable job and fulfill the thirst of actualizing this intellectual capacity. The dependence of modern economy on fossils, industry and agriculture with a high share of monotonous labor makes this side effect a real threat for an excessive number of people with highly developed intellect. The possible solution would be to use additionally some kind of mood or discipline enhancement. However, let us assume that at the current stage of development in biomedical science, the discipline enhancement would be incompatible with the mental enhancement. They neutralize each other and an enhancement package, which includes both, would simply be a waste of money.
Clearly, it is the non-ideal circumstances, such as current development of science, society and the state of human nature which affect an evaluation of the enhancement, even if we think that it objectively moves us toward the state of the ideal. The presence of good genes/nature themselves does not warrant the positive and useful realization of them in the future. In the aforementioned example with IQ the pursuit of well-being needs, as a prerequisite, an enhanced nature, which is available and a certain social structure, which is not available. Thus, the implementation of this enhancement ostensibly brings us closer to the ideal, but in reality it does not contribute to the well-being and consequently it cannot be evaluated as a positive intervention. It seems legitimate to argue instead that our evaluation of these would-be biomedical interventions should be practical rather than ideal.

When we face a choice, such as that described above, the evaluation of a possible human enhancement with an ideal yardstick will not suffice for our decision-making. Following Amartya Sen and his critique of ideal theory I affirm that the ideal approach is not only insufficient for assessing a current state, but it is also unnecessary (Sen 2011, 15-16; 87-112). Objectively, the similarity of methodology for talking about ideal and non-ideal scenarios in biomedical enhancement and in political philosophy seems to be obvious. Therefore, the application of some conclusions from the developed arguments about ideal/non-ideal theories in political philosophy to the field of bioethics does not look as something artificial and inconsistent. The irrelevance of an ideal for assessment of a current state is revealed by means of attention to present physical deficiencies (injustices in the case of social philosophy). A dissatisfaction with a particular trait of a human and a desire to enhance it does not necessarily signal our longing for an ideal state. Our dissatisfaction with an unfavorable state does not come from deprivation of an ideal state. Injustices (in social philosophy) and disadvantages (in biomedical science) have their own ontological ground, so to speak; they exist as unsatisfactory conditions regardless of whether we have an ideal in mind or not. Biological disadvantages or deficiencies can be defined as privations, which is an essential principle of change, and any enhanced state is possible only because the state of deficiency or disadvantage has a potential for a change. Thus, we need to have a concept of the opposite to privation, but this opposite is not necessarily an ideal state.

The unnecessariness of an ideal, according to Sen, comes from its inability to decide which of the two alternatives is preferable. In his famous example, Mona Lisa as a tentatively best masterpiece is of no help in deciding who is better Van Gogh or Picasso (Sen 2011, 101). Instead of a transcendental ideal Sen advocates a comparative approach according to which an evaluation of two alternatives can be made “without
a prior identification of a supreme alternative” (Sen 2011, 102). The two alternatives which qualify for independent comparison would be two states of justice for Sen. If applied to biomedical enhancement two alternatives for an independent comparison can be either an enhanced state vs the status quo ante, or an enhanced state vs some kind of other state. If one of the alternatives is status quo ante then evaluation immediately becomes backward-looking. The previously described shortsightedness of the backward-looking approach forces me to claim that the second pair of alternatives is the only available option.

So far, I have not stated explicitly what would be an appropriate way for evaluating enhancement. There has been enough said about what it would not be though. To be sure, the assessment of biomedical enhancement should not start with figuring out what the ideal of man is; it should not end with figuring out this question either. The starting point consists in figuring out a state liable to amelioration. This, I have argued, is a privation of a certain state, and this privation is an essential principle of transition to an enhancement state. At the same time, this evaluating approach is forward-looking in the sense that it is comparative and presupposes some better state, which is ahead of status quo ante.

The idea of a better state as a point of reference for evaluation of human enhancement has three crucial features, which help to qualify it as non-ideal. 1) It concedes multiplicity. Plurality of values does not allow sticking to one unified standard of better. Genetic augmentation of breasts or improvement of hearing is a desirable and legitimate enhancement for some but not necessarily for all. 2) The better state does not regulate all possible enhancements, but simply allows evaluating particular solutions based on shared values, which are necessary for a better life. 3) It is flexible, i.e., it can, and if need be it has to be revised depending on: a) current state of human nature; b) social factors; c) technological capabilities. Enhancement is a type of adaptation like evolution, which is not teleological, but is navigated by a better solution for a more convenient living and coping with present circumstances. For example, admittedly the constitution of our backbone is not suitable for the amount of everyday sitting we experience nowadays. Our nature definitely adjusts to these changes in lifestyle, but those natural processes are extremely slow. Up until our nature accommodates to these new challenges of ‘environment’ a large number of generations will suffer from pain, which is unbearable for many, due to disc herniation, pinched nerves, pressures on the spinal cord etc., simply because of sitting for longer time than our body can ‘afford’. If we imagine that, a would-be enhancement can do the work of evolution, we will have an attuned nature more efficiently and faster. This
will be a definite improvement of our life, but it will not be an ideal one. This desired life with no pain in the back is not an ideal state, because its effectiveness depends on a number of contingent factors. Namely, for instance the molecular level of human organism requires our body to move, standing and moving activate the molecular processes in muscles and the cellular system: blood sugar, triglyceride, cholesterol, etc. Movement is something what works as a life pump for those processes; lack of movement results in a gradual but imminent attenuation, followed by deterioration and finally complete failure. Thus if we imagine that an enhancement package solves the problem of pain in the back but neglects these aforementioned problems, our better state with regard to this problem should be revised and has to take into account this peculiarity of human nature. Life without pain in the back is better and desirable but this enhancement cannot be evaluated positively if it does not address the influence of sitting on molecular processes.

In my example about enhancement packages which either include mood enhancement or intellectual enhancement, I tried to show that, given some social constraints, a mass non-cognitive enhancement can seem more appealing than improvement of the mental state. The better state, which would take into account that social constraints can give us a goal that is able to provide normative guidance for enhancement policy, which needs to be effective and technically possible.

Therefore biomedical enhancement as a pursuit of improvement is a phased process, whose length, intensity and trajectory are dictated by the three mentioned above factors (human nature, available technologies, and social structure). Upon completion of a particular phase the length, intensity and trajectory are subject to revision depending on the understanding of the next step of a better life and the ways it can be pursued. This better state resembles the concept of a desired functioning, and it should not be confused with some ideal, because this desired state is partly formed by our intentions and understanding of a good life, but at the same time a large number of contingent factors has a great deal of impact on it.

One of the contingent factors, which seems to me crucial for our inability to assess biomedical enhancement by means of ideals is the nature of human beings. Functioning of the organism from the point of view of evolutionary biology is not something optimal or harmonious but it is a result of a long developmental process, which has no notion of human well-being in mind. Despite the fact that desired functioning (the better state) is supposed to evaluate the enhanced state, as some product of creativity, which was designed to make us better off, human organism, even after the application of some enhancement is not a balanced or completed whole; we
are dynamic creatures who still evolve and enhancement is a part of this evolvement. Prior to any enhancement human nature is a “tentative, changing, perishing, cobbled-together ad hoc solutions to transient design problems, within blithe disregard for human well-being” (Buchanan 2016, 2). However, it does not mean that after several interventions by means of biomedical enhancement our nature is no longer a tentative solution to transient problems. It is exactly the opposite; gradual enhancement is nothing else but a number of tentative solutions on the way to a good life. It is worth noting that these tentative solutions are not arbitrary, but they are situational, and the following trajectory of the future solutions is to be revised along with every step of enhancement.

References
Joseph M. Dunne
The University of Michigan–Dearborn

Biography
Joseph M. Dunne is a Lecturer in Philosophy at The University of Michigan–Dearborn. He holds a PhD in Philosophy from Wayne State University and his research interests include moral philosophy, law and religion, bioethics, and the philosophy of religion.

Publication Details

Citation
Content Neutrality: A Defense

Joseph M. Dunne

Abstract
To date, both the United States federal government and twenty-one individual states have passed Religious Freedom Restoration Acts that aim to protect religious persons from having their sincere beliefs substantially burdened by governmental interests. RFRAs accomplish this by offering a three-pronged exemption test for religious objectors that is satisfied only when (1) an objector has a sincere belief that is being substantially burdened; (2) the government has a very good reason (e.g., health or safety) to interfere; and (3) there is a reasonable alternative to serve the compelling interest. Legal balancing tests like those found in RFRA are content neutral insofar as they sideline the belief-content of conscientious objections as irrelevant when determining the permissibility of granting legal accommodations. However, some theorists worry that this legal picture may be backward: perhaps balancing tests should be content non-neutral given the usual features of conscientious objections. For example, Yossi Nehushtan contends that, contrary to their typical codification, religious conscience beliefs seem undeserving of special legal accommodations because they possess uniquely strong empirical and theoretical ties to intolerance. Thus, the illiberally intolerant content of these conscientious objections might actually give the state a reason to refuse to grant legal exemptions. In this paper, I offer a cursory defense of content neutrality with respect to balancing tests like those found in RFRA. To begin, I outline Nehushtan’s argument for content non-neutrality. The cornerstone of his argument is that illiberal intolerance is intolerable such that conscientious objections that are based upon illiberally intolerant values provide the state with strong, normally prevailing reason not to grant an exemption. I argue that, even when the illiberally intolerant content of one’s conscience constitutes a weighty and relevant factor in determining the permissibility of granting a legal exemption, there remain significant problems. It is difficult, for example, to determine which views are illiberally intolerant and difficult to say whether illiberally intolerant views can effectively serve as the principled demarcating line in balancing tests. To conclude, I offer several cursory arguments in favor of adopting content-neutral approaches without necessarily making a comprehensive case. By drawing on the work of Amy Sepinwall, Nadia Sawicki, and Nathan Chapman, I show that content-neutral approaches can help to safeguard robust protections for conscience by permitting atypical exercises of conscience, protect minority thoughts and practices from being coercively supplanted by majoritarian understandings of morality, appropriately maintain the skepticism and humility that we owe each another as compatriots in a pluralistic society, and allow the kind of justifiable civil disobedience that has an important place in political history among other things.

Keywords
Conscientious Objection(s), Legal Exemption(s), Content Neutral Balancing Test(s), Religious Freedom Restoration Act(s), Conscience

I. Introduction
To date, both the United States federal government and twenty-one individual states have passed Religious Freedom Restoration Acts that aim to protect religious
persons from having their sincere beliefs substantially burdened by governmental interests. RFRAs accomplish this by offering a three-pronged exemption test for religious objectors that is satisfied only when (1) an objector has a sincere belief that is being substantially burdened; (2) the government has a very good reason (e.g., health or safety) to interfere; and (3) there is a reasonable alternative to serve the compelling interest. Legal balancing tests like those found in RFRA are content neutral insofar as they sideline the belief-content of conscientious objections as irrelevant when determining the permissibility of granting legal accommodations. However, some theorists worry that this legal picture may be backward: perhaps balancing tests should be content non-neutral given the usual features of conscientious objections. For example, Yossi Nehushtan contends that, contrary to their typical codification, religious conscience beliefs seem undeserving of special legal accommodations because they possess uniquely strong empirical and theoretical ties to intolerance. Thus, the illiberally intolerant content of these conscientious objections might actually give the state a reason to refuse to grant legal exemptions.

In this paper, I offer a cursory defense of content neutrality with respect to balancing tests like those found in RFRA. To begin, I outline Nehushtan’s argument for content non-neutrality. The cornerstone of his argument is that illiberal intolerance is intolerable such that conscientious objections that are based upon illiberally intolerant values provide the state with strong, normally prevailing reason not to grant an exemption. I argue that, even when the illiberally intolerant content of one’s conscience constitutes a weighty and relevant factor in determining the permissibility of granting a legal exemption, there remain significant problems. It is difficult, for example, to determine which views are illiberally intolerant and difficult to say whether illiberally intolerant views can effectively serve as the principled demarcating line in balancing tests. To conclude, I offer several cursory arguments in favor of adopting content-neutral approaches without necessarily making a comprehensive case. By drawing on the work of Amy Sepinwall, Nadia Sawicki, and Nathan Chapman, I show that content-neutral approaches can help to safeguard robust protections for conscience by permitting atypical exercises of conscience, protect minority thoughts and practices from being coercively supplanted by majoritarian understandings of morality, appropriately maintain the skepticism and humility that we owe each another as compatriots in a pluralistic society, and allow the kind of justifiable civil disobedience that has an important place in political history among other things.
II. An Argument for Content Non-Neutrality

Yossi Nehushtan offers a noteworthy argument in favor of content non-neutrality with respect to legal exemption tests. First, he contends that the limit of liberal tolerance should be illiberal intolerance – i.e., that “illiberal intolerance should almost never be tolerated by the tolerant-liberal state” (Nehushtan 2018, 198). Second, he contends that there are “meaningful and special” empirical and theoretical links between religions like Christianity, Islam, and Judaism – i.e., the Abrahamic religions – and intolerance (Nehushtan 2018, 198). As a result of these “special” links, he thinks that Abrahamic religionists are “more likely to be intolerant and act intolerantly than non-religious people” (Nehushtan 2018, 198). Under the assumption that his first two claims are true, he concludes that when they’re taken together, these claims “may provide a reason for the tolerant-liberal state not to tolerate religion as such” (Nehushtan 2018, 198). At the very least Nehushtan thinks that these two claims “provide a reason for the tolerant-liberal state not to tolerate religious claims for accommodation or religious claims for exemptions from legal rules” (Nehushtan 2018, 198). This means that Nehushtan would ultimately answer the “Is religion special?” question in the affirmative, but with a catch: since religious beliefs more broadly – which include religious conscience beliefs – are meaningfully and specially linked to illiberal intolerance, we should afford them a special, negative status before the law.

This argument suffers from a few mistakes worth highlighting. First, in order for the two central claims of the argument to be congruent, Nehushtan would need to argue that there are meaningful and special empirical and theoretical links between the Abrahamic religions and illiberal intolerance – not just any kind of intolerance. After all, Nehushtan’s argument claims that at least some forms of intolerance are tolerable – namely, intolerance of illiberal views or values. Second, while the Abrahamic religions may possess a meaningful relationship to illiberal intolerance, it is not obvious that they possess a special or unique relationship to such intolerance. Both secular and religious conscience beliefs, for example, plausibly seem to share a similar in-principle and in-practice link to illiberal intolerance such that special legal treatment for either form of conscience might be unjustified.

To see how both secular and religious conscience beliefs may share a similar in-principle link to illiberal intolerance, consider the example of suicide bombers offered by social-psychologist Jonathan Haidt:

To take one example, religion does not seem to be the cause of suicide bombing. According to Robert Pape, who has created a
database of every suicide terrorist attack in the last hundred years, suicide bombing is a nationalist response to military occupation by a culturally alien democratic power. It’s a response to boots and tanks on the ground – never to bombs dropped from the air. It’s a response to contamination of the sacred homeland. (Imagine a fist punched into a beehive, and left in for a long time.) Most military occupations don’t lead to suicide bombings. There has to be an ideology in place that can rally young men to martyr themselves for a greater cause. The ideology can be secular (as was the case with the Marxist-Leninist Tamil Tigers of Sri Lanka) or it can be religious (as was the case with the Shiite Muslims who first demonstrated that suicide bombing works, driving the United States out of Lebanon in 1983). Anything that binds people together into a moral matrix that glorifies the in-group while at the same time demonizing another group can lead to moralistic killing, and many religions are well suited for that task. Religion is therefore often an accessory to atrocity, rather than the driving force of the atrocity. (Haidt 2013, 312)

According to Haidt (2013), illiberal intolerances – like those manifested in suicide bombings – are not the unique products of religion per se, but instead seem to be the products of ideologies that “bind people together into a moral matrix that glorifies the in-group while at the same time demonizing another group.” While these ideologies can be either secular or religious, what serves as the primary “driving force of the atrocity” is being deeply embedded within the above sort of moral matrix that glorifies the in-group and demonizes the out-group. Thus, the underlying feature that seems to drive such illiberal intolerance is not, at least in principle, unique to religious conscience beliefs.

Moreover, it seems doubtful that illiberally intolerant ideologies are, in practice, uniquely religious. For example, one of the main findings of the BBC’s war audit in 2004 was that the “overwhelming majority of wars and the overwhelming majority of the victims of such wars cannot be classified primarily according to religious causes or religious beliefs” (Austin, Kranock, and Oommen 2004, 17). There are, of course, obvious examples where religious communities have been targeted because of their faith however. Several of the most heinous atrocities against communities of faith in particular were perpetrated “by the three most vicious and blood-thirsty regimes ever to hold power: Stalin’s Russia, Mao’s China and Hitler’s Germany” (Austin, Kranock,
and Oommen 2004, 17). These specific examples among others provide “a fairly strong indication” that atheistic or secular states are not “less prone to war or large scale violence” – but perhaps even more prone (Austin, Kranock, and Oommen 2004, 35-36). Like Haidt, the audit argues that the common thread “linking the disposition to war of religious and atheistic states is absolutism: the more absolutist the state, the more likely it is to go to war” (Austin, Kranock, and Oommen 2004, 35-36). Thus, they conclude that a “genuinely secular (atheistic) state may be less inclined to go to war than a state in which religion is very prominent, only as long as the secular state is one which is not pursuing a millenarian or totalitarian ideology (such as Communism or Nazism) and as long as the state is one in which pluralism and tolerance of diversity are the norm” (Austin, Kranock, and Oommen 2004, 35-36; emphasis added).

Moving forward, we can offer some friendly amendments to Nehushtan’s argument given the above claims. First, the foundational claim of his argument can remain the same: illiberal intolerance should not be tolerated. But second, it seems like we can reasonably deny that there are special or unique in-principle and in-practice links between the Abrahamic religions and illiberal intolerance. Our examples seem to indicate that secular conscience can be in principle and in practice just as illiberally intolerant as the Abrahamic religions can be. Accordingly, we can deny that religious beliefs should be afforded a special, negative status before the law. Instead, any conscience belief – secular or religious – that relies on illiberally intolerant values can be afforded that special, negative status before the law. In short, the illiberally intolerant nature of any conscience belief may provide a reason for the tolerant-liberal state not to tolerate that claim for accommodation or for exemptions from legal rules.

With these amendments at hand, what sort of problems now remain with the view that the illiberally intolerant content of anyone’s conscience may provide a reason not to grant a legal exemption? There seem to be at least three. First and foremost, there is a Conceptual Problem: it is not clear that illiberally intolerant views adequately serve as the principled, demarcating line in balancing tests for legal exemptions. In other words, it is not clear that, in principle, every illiberally intolerant conscience belief should fail to receive legal accommodations. To illustrate, just consider the cases involving the ministerial exception.2 In those cases, churches were exempted from anti-discrimination

---

1. They continue: “Atheist governments in the USSR, China and Russia were…the biggest perpetrators of mass violence that the world has ever seen, with both governments individually responsible for many more deaths than the Nazi regime of Adolf Hitler” (Austin, Kranock, and Oommen 2004, 35-36).

2. See, for example, Hosanna-Tabor Evangelical Lutheran Church & Sch. v. E.E.O.C., 565 U.S. 171 (2012).
Dunne

laws and the application of these civil rights when hiring employees, leaving matters of internal church governance up to the church in question. Such cases would disallow the United States government from, say, forcing a Catholic Church to hire a female priest. The conscientious content in these cases is, on its face, illiberally intolerant – yet it is not clear that legal exemptions should be withheld in these cases. Indeed, extending legal exemptions to churches in matters of internal church governance may be the morally right thing to do for several reasons not least of which is maintaining the appropriate and principled division between Church and State.

Second, there is an Epistemic Problem for courts that try and determine whether the content of some conscience belief is, in fact, illiberally intolerant or not. Though the content of some conscience beliefs is obviously illiberally intolerant – e.g., racist beliefs – the content of other conscience beliefs is not so obviously illiberally intolerant – e.g., anti-abortion beliefs. Asking courts and judges to be the final adjudicator of such contested matters of moral philosophy and epistemology seems, at best, highly suspect.

Third, and finally, let’s suppose that the first two problems are surmountable and that the illiberally intolerant content of a conscience belief can be considered a “relevant” and “strong or weighty reason” for not tolerating that claim (Nehushtan 2018, 200). Given these assumptions, the state would then have a good, though “not necessarily conclusive reason” not to grant conscientious exemptions – “even in cases where the claims to be granted exemptions are not based directly on intolerant values” (Nehushtan 2018, 200). The primary concern with this view is this: even when we grant that the content of one’s conscience constitutes a relevant and weighty factor in deciding whether or not to grant an exemption, just how weighty and relevant that reason is or should be remains unhelpfully unclear – especially when considered among the balance of competing reasons for granting or not granting an exemption. We can call this the Weightiness Problem. Though Nehushtan contends that the illiberally intolerant content of some conscience belief should be a relevant and weighty factor, he nevertheless concedes that this factor “should not, however, be an overriding or conclusive reason for not tolerating such claims” (Nehushtan 2018, 200). Because “there are always reasons for tolerating [illiberally intolerant] claims for accommodation or exemption,” those reasons may “override the reason against tolerating such claims and may therefore justify accommodation or granting exemptions” (Nehushtan 2018, 200). In the end, then, how helpful and how much does a content non-neutral approach really accomplish if content-based reasons are just one set of relevant and weighty though non-overriding reasons within the balance of reasons? Just how
these reasons might factor into the exemption calculus is unclear and uncertain under content non-neutral approaches like Nehushtan’s – and this seems to be as relevant and weighty of a reason as any not to endorse such views.

III. Some Arguments for Content Neutrality

In what remains, I offer several cursory arguments in favor of adopting content-neutral approaches without necessarily making a comprehensive case. By drawing on the work of Amy Sepinwall, Nadia Sawicki, and Nathan Chapman, I highlight some of the virtues of adopting a content-neutral approach over a content non-neutral approach.

A. Sepinwall’s Third-Party Harm Argument

Amy Sepinwall offers a defense of content neutrality that arises in the context of addressing a separate though related question: when is a claim of moral complicity\^3 compelling enough to warrant an accommodation? Her answer is that conscientious objections should be evaluated under a sort of content-neutral RFRA regime, but that “a separate, additional set of considerations must be brought to bear – namely, considerations tracking the interests of third parties” (Sepinwall 2015, 1974).\^4 So, she suggests adding an additional feature to the content-neutral RFRA test for legal exemption: roughly, third-party costs “exceeding some threshold amount should be found untenable” (Sepinwall 2015, 1974). Exemptions would be denied, therefore, when these excessive and unacceptable third-party costs would otherwise result (Sepinwall 2015, 1974). Specifying exactly where this threshold might be on a cost spectrum is, Sepinwall concedes, not always clear and “a matter for democratic deliberation” (Sepinwall 2015, 1974).

Sepinwall adds and emphasizes this more extrinsic feature to the test of exemption demarcation because she believes that there are good reasons for courts to abstain from evaluating the intrinsic content of any given conscience belief. More specifically, she thinks emphasizing this additional feature will ameliorate the Epistemic Problem: since courts are epistemically limited when evaluating the more intrinsic features of

---


4. Arguably, this feature may already exist in the second, “governmental interest” prong of the RFRA test. I will set aside that discussion for now and assume that Sepinwall’s feature is not already implicit or embedded within the RFRA test.
a conscience belief, they should instead focus on the extrinsic features of requested accommodations when determining whether granting a legal exemption is warranted or not. In fact, Sepinwall argues that moral beliefs in general should largely enjoy total deference by courts for this very reason. She thinks that a position of total deference to an objector’s moral beliefs seems to appropriately maintain the “skepticism and humility that we owe one another as compatriots in a pluralistic society” (Sepinwall 2015, 1927). And given the reality of moral and religious pluralism, “we are often without a capacity for certitude…in moral and religious matters…that would allow us to discern truth and falsity” (Sepinwall 2015, 1927). So, it seems like an attitude of “moral deference on the part of the state” is most appropriate (Sepinwall 2015, 1927).

At this point, we should wonder: under content-neutral approaches, what will become of particular moral beliefs that are obviously wrong and illiberally intolerant—like racism? Should certain limits on moral deference and content-neutrality be appropriated so as to ensure that courts, as state actors, are not compelled to treat highly problematic moral beliefs as on par with other moral beliefs? We can offer a few responses to this Problem of Bad Views via Sepinwall. First, she contends that “according deference to a claim that denigrates another group is not the same as endorsing that claim” (Sepinwall 2015, 1928). Though a court should treat the moral belief with deference and neutrality, Sepinwall thinks that courts have a responsibility to clearly articulate that such a moral view “flies in the face of our most fundamental constitutional values…[thereby serving, say,) religious freedom but also [speaking] in favor of the notion of equal respect that underpins our constitutional regime” (Sepinwall 2015, 1928). Even Nehushtan agrees that allowing the authorities to grant an exemption yet “condemn their values or behavior” may be a justifiable way of “not tolerating intolerant conscientious objectors” (Nehushtan 2018, 199-200).

5. Interestingly, though Sepinwall grants much deferential latitude for religious claims of complicity, she nevertheless thinks that “we should expect that claims seeking religious exemptions from antidiscrimination laws would typically fail” because the third parties whose interests are relevant in these cases are those who are “immediately denied service or employment by the religious objector” (Sepinwall 2015, 1978). Moreover, she thinks that “all members of the group facing discrimination can claim an expressive injury from the discrimination” and that “other historically oppressed groups can claim that an exemption threatens them with an injury, too” (Sepinwall 2015, 1978). A state that grants an exemption to the religious objector in this case runs the risk of failing to “take seriously the great evil of discrimination” and undermining “the sense of security and respect that a decent state should confer on all its citizens” (Sepinwall 2015, 1978).
Second, deferring to and remaining neutral toward a moral claim doesn’t necessarily commit a court to issuing an exemption: the objector’s assertion must still be weighed against governmental interests. Hence, the entitlements that citizens enjoy would be defeasible. It may turn out, for example, that in some instances, “the government will invoke its compelling interest in the eradication of, say, racism, and it will wield that interest to defeat the bid for an exemption” (Sepinwall 2015, 1928). Remember too that the moral belief must also be weighed by the court against the interests of third parties – at least according to Sepinwall’s revised test. She writes:

Third parties will presumably be able to marshal arguments that acceding to the believer’s hateful claim inflicts a grave injury on them—one so grave that the court should find it dispositive. But even if third parties choose not to become too vexed about the believer’s claim, the state must, again in its capacity as a defender of our constitutional regime, add to its arguments about the compelling interests underpinning the challenged legal requirement a statement decrying the challenge because it deviates from our most cherished constitutional values. (Sepinwall 2015, 1929)

As previously noted, Sepinwall thinks that adding and emphasizing this third-party harm feature will ameliorate the Epistemic Problem. But her feature cannot, however, fully eradicate this problem: though determining resultant harm may be an easier and more suitable task for the courts, they will nevertheless struggle, at times, to determine when an excessive and unacceptable third-party cost has resulted. Consider, for example, the plethora of cases involving religious wedding vendors and servicers conscientiously refusing to offer their services for same-sex weddings. You might think, as Sepinwall seems to, that their conscientious objection amounts to a harm sufficient for blocking an exemption from antidiscrimination law – and sufficient for the reasons articulated above (Sepinwall 2015, 1928). But you might think, as Eugene Volokh and Dale Carpenter, that no considerable harm is committed against same-sex couples when, say, a wedding photographer turns them away. They write:

[D]iscrimination by these narrow categories of expressive commercial actors is much less damaging and restrictive than other forms of discrimination. Employment discrimination can jeopardize a person’s livelihood. Discrimination in education can affect a person’s future, as can discrimination in housing—especially when housing is scarce.
in the safe parts of town with good schools. Discrimination in many places of public accommodation has been historically pervasive, to the point that mixed-race groups might have been unable to find any suitable hotel or restaurant. But protecting the First Amendment rights of writers, singers, and photographers would come at comparatively little cost to those denied such inherently expressive and personal services by specific providers. Of course, when a photographer tells a couple that she does not want to photograph their commitment ceremony, the couple may understandably be offended by this rejection. But the First Amendment does not treat avoiding offense as a sufficient interest to justify restricting or compelling speech. (Volokh and Carpenter 2013, 19-20)

As Volokh and Carpenter note, it seems doubtful that the rejection of an expressive business rises to the level of an injury so grave and excessive that the courts should find it dispositive. Of course, there are cases where a photographer’s rejection could arguably harm a same-sex couple in an excessive way – namely, by withholding services that are not otherwise available to them. For example, Kevin Vallier argues that a genuine and noteworthy harm was committed to same-sex couples when Kim Davis – the county clerk from Rowan County, Kentucky – refused (and refused to allow her willing deputies) to certify marriage certificates for same-sex couples that were legally available to them (Vallier 2015). However, by refusing to use their expressive talents to photograph a same-sex wedding, a photographer will likely not withhold services or interests that are not otherwise available to the couple in the way that Davis did. As Kevin Vallier explains, the actions of a religious wedding servicer “may offend, but they do not harm the couple on any sensible understanding of harm” (Valleir 2015).

Just to reiterate Sepinwall’s earlier claim: given this new feature, exemptions would be denied when excessive and unacceptable third-party costs would otherwise result. But specifying exactly where this threshold might be on a cost spectrum is not always clear. In cases involving racism, the resultant harm would be so clearly excessive and unacceptable that requested accommodations would absolutely fail. But in cases involving religious wedding servicers and discrimination against same-sex couples, the

6. Vallier adds: “As John Stuart Mill taught us, any free society must distinguish between harms and offenses. We have trouble with that distinction, but it is critical for freedom” (Vallier 2015).
resultant harm may not be comparably as potent. John Corvino offers some helpful insight as to why the resultant harm of these two cases may be different after all:

When civil rights laws were passed, discrimination against blacks was pervasive, state-sponsored, and socially intractable. Pervasive, meaning that there weren’t scores of other photographers clamoring for their business. State-sponsored, meaning that segregation was not merely permitted but in fact legally enforced, even in basic public accommodations and services. Socially intractable, meaning that without higher-level legal intervention, the situation was unlikely to improve. To treat the lesbian couple’s situation as identical – and thus as obviously deserving of the same legal remedy – is to minimize our racist past and exaggerate L.G.B.T.–rights opponents’ current strength. (Corvino 2015)

In sum, Sepinwall argues that courts should afford a great deal of deference and neutrality when facing a conscientious objection. After all, “pluralism demands respect for religious differences, but that respect goes both ways: it entails that we must be open to many claims of conscience, but we must also ensure that these claims do not unduly or disproportionately interfere with the interests of discrete third parties (Sepinwall 2015, 1972). However, the deferential conclusion sketched here does not necessarily entail that an objector is always entitled to an exemption. Even though the courts “should in general treat as true the …adherent’s claim…, they must still consider whether acceding to a request for an accommodation would impose undue burdens on third parties” (Sepinwall 2015, 1910).

B. Sawicki’s Respect for Conscience Argument

Sawicki’s defense of content neutrality begins with her claim that true respect for any conscience belief demands a consistent, coherent, and repeatable mechanism for legal accommodation, even if that test is open-ended and results in uncertainty at the outermost margins (Sawicki 2012, 1395). As such, she thinks that the most promising legal mechanism for meeting these demands is the kind of content-neutral balancing test we see in RFRA (Sawicki 2012, 1396). The alternative is to include content as a relevant feature within the mechanism for legal accommodation – something that Sawicki fears may yield less consistent results, produce less coherent justifications, and may, as a result, lose repeatability as a mechanism over time. Of course, she grants that balancing approaches are subject to some criticism – e.g., they may risk being used as
a proxy for judgments based on majoritarian values – but she nevertheless contends that the alternative to establishing a content-neutral guiding principle is to flatly “abandon the promise of freedom of conscience and concede that American society considers exercises of personal conscience to be valuable only to the extent that they align with widely accepted moral principles” (Sawicki 2012, 1396). Such an alternative, Sawicki argues “would undermine the foundational purpose of legal accommodation of conscientious belief, which is to protect individuals from oppressive majoritarian understanding of morality” (Sawicki 2012, 1396). Thus, she contends that “only a content-neutral approach…is consistent with the principles of a pluralistic society that respects the inherent value of conscience” (Sawicki 2012, 1445).

Moreover, Sawicki worries that the pragmatic implications of failing to adequately respect conscience beliefs and protect them from oppressive majoritarian understandings of morality may be socially unworkable as well. She notes that “we begin to understand the pragmatic argument for legal accommodation of conscience” only when we “accept that an individual may be driven to act in accordance with her strong conscientious beliefs regardless of the consequences” (Sawicki 2012, 1404). Indeed, those in the grip of conscience often feel the compulsion of conscience so deeply that they say that they could not live with themselves if they were to violate their conscience (Sorabji 2014, 36). So, Sawicki’s argument for accommodating conscience beliefs maintains “that a civil society is unlikely to function effectively if it chooses to punish conscientious objectors” (2012, 1405). She writes:

Punishment in a civil society can serve many purposes, among them retribution, deterrence, and reform. However, punishing a person for acting in accordance with her conscience rarely serves these purposes effectively. First, because it is impossible to coerce belief against someone’s wishes, punishment of conduct motivated by conscientious belief is unlikely to result in reform or rehabilitation. On a similar note, the threat of punishment is unlikely to deter those acting on the basis of conscientious conviction. That being said, some conscientious objectors (those who are more are susceptible to the threat of punishment, perhaps because their moral commitments are not as firm) may be swayed by the threat of legal penalty, and may choose to comply with the law despite their conscience’s voice to the contrary. What is the likely result for these persons, we might ask? “Deterrence of those who lack the will to act on their
convictions exacts a terrible price. Their feeling that they have yielded to compulsion and violated their most deeply held beliefs and principles may involve profound resentment and loss of self-respect.” Finally, when it comes to retribution, many believe that the retributive purposes of punishment are poorly served by punishing individuals who act on the basis of moral compulsion rather than self-interest or impulsiveness. (Sawicki 2012, 1404-05)

Overall, Sawicki’s pragmatic argument for accommodating conscience beliefs recognizes the punishment concerns raised above and responds accordingly. Of course, you may worry that “granting conscience-based exemptions from legal obligations as a matter of course may wreak havoc on the state’s ability to maintain order” (Sawicki 2012, 1405). But her point is that “the same can be said of a state that rejects claims of conscience altogether” (Sawicki 2012, 1405). Because a society cannot adopt laws that perfectly align with all of its citizens’ beliefs, Sawicki contends that a “society must order itself in such a way as to mediate between the interests of social order and the interests of citizens who might feel disenfranchised by laws that violate their conscientious beliefs” (Sawicki 2012, 1405-06). And perhaps the only pragmatic and workable way to accomplish this, as Sawicki thinks, is to offer adequate, content-neutral accommodations for conscience beliefs like those discussed above.

C. Chapman’s Anti-Tyranny Argument

Finally, Nathan Chapman also offers several reasons for why we ought to adopt broad, content-neutral, defeasible entitlements for conscience beliefs. Perhaps most importantly, he argues that granting such entitlements for conscience may help to protect against tyranny insofar as “protecting conscience undermines the totalization of morality by the government” (Chapman 2013, 1494). Like Sawicki, Chapman contends that content-neutral entitlements in particular can limit the “government’s pretensions to absolute moral authority” by permitting “nonconformist moral thought that aims to undermine moral tyranny” (Chapman 2013, 1499). These protections may be especially important for minority thoughts and practices that wish to persist against the conclusions promoted by the majority. These protections may further allow for “majority decisions to be provisional,” allow the “persuasiveness of minority speech to be aided by the persuasiveness of minority action,” and allow for those who disagree with prevailing norms to “prolong internal and national dialogues over contested moral issues” (Chapman 2013, 1500). Additionally, the ability to contribute to the kind
of justifiable civil disobedience that has “an important place in political history and theory” – the kind that can be “particularly effective at jarring a morally apathetic society into taking notice and making important changes” – will be increasingly secured as well (Chapman 2013, 1499).

Of course, the virtues of offering broad, defeasible, and content neutral entitlements are best displayed when the moral views of the government or majority are obviously bad. But it seems like such entitlements should be offered during the good times and the bad – at the very least – in order to uphold the suitable level of consistency and toleration. After all, tolerance requires, as Brian Leiter suggests, not merely being indifferent toward some group, but actively putting up with the perceivably wrong, mistaken, or undesirable beliefs and actions of that group (Leiter 2013, 8). A tolerant attitude toward conscientious objectors, therefore, would seemingly allow them the opportunity to be wrong by permitting a legal exemption in some cases. Otherwise, we may run the risk of a kind of “totalization of morality” on the part of the government stressed above. Moreover, we should remember that every permitted accommodation requires at least some degree of tolerance insofar as a conscientious objection is, by nature, requesting exemption from a law that is otherwise taken to be correct or justified. As Yossi Nehustan notes, the very act of granting a conscientious exemption “presupposes that the state does not share the conscientious objector’s values or his way of balancing between values” (Nehushtan 2018, 130). Indeed, the state “believes it would be unbearable and indeed intolerable if everyone shared the objector’s kind of conscience and reasoning” (Nehushtan 2018, 130). So, tolerance is required to grant any legal exemption to any kind of conscientious objector – even in cases when the law is wrong and the conscientious objector is right.

References


Vallier, Kevin. 2015. “Bakers and Clerks: 5 Differences.”
The Ethics of Knowing the Score: Recommendations for Improving Boxing’s 10-Point Must System

John Scott Gray
Ferris State University

Brian R. Russ
Indiana University - Purdue University of Columbus

Biographies
John Scott Gray is a Professor of Philosophy and Distinguished Teacher at Ferris State University in Big Rapids, MI. His research interests focus on areas of applied philosophy, including the Philosophy of Sports, Sex and Love, and Popular Culture. Dr. Gray co-authored Introduction to Popular Culture: Theories, Applications, and Global Perspectives (2013). He lives in Canadian Lakes, MI with his wife Jo and son Oscar.

Brian R. Russ is an Assistant Professor of Mental Health Counseling at Indiana University – Purdue University of Columbus in Columbus, IN. His research interests include sports counseling and adolescent mental health. Dr. Russ lives in Columbus with his wife and three children, and he is a lifelong boxing fan.

Publication Details

Citation
The Ethics of Knowing the Score: Recommendations for Improving Boxing’s 10-Point Must System

John Scott Gray and Brian R. Russ

Abstract
The sport of boxing has been a mainstay in urban settings since the late 19th century. Although there have been changes in the sport to enhance safety and justice, the 10-Point Must scoring system continues to be problematic. Given the amount of money at stake at the highest levels of professional boxing, as well as the health of the fighters themselves, this paper will argue that professional fighting has a moral obligation to bring their judging system into the 21st century. Ethical principles such as justice, autonomy and nonmaleficence demand this revision, as the current scoring procedures are systematically problematic. In this article, recommendations are made to improve the 10-Point Must System, including an argument for open scoring.

Keywords
Boxing, Ethics, Open Scoring, Cognitive Bias

The Scope of the Problem
While Joyce Carol Oates might rightly describe boxing as a sport both in crisis and “a sport of crisis,” professional boxing has long been ridiculed for poor decisions stemming from the wide range of judges’ scores that can result from a fight going the distance (2006, viii). A recent pay-per-view fight between Gennady Golovkin and Canelo Alvarez was called a draw, after one judge had the fight won by Golovkin 115-113 (seven rounds to five) and another judge had the same fight 118-110 (ten rounds to two) for Alvarez, amounting to an astonishing five round difference in scoring.

Another example of a scoring discrepancy can be found in the 2004 Courtney Burton vs. Emanuel Augustus match, which prompted an investigation. ESPN play-by-play man Joe Tessitore during the fight even warned those watching at home, “the Michigan judges have been known to come up with some curious results.” Augustus landed 71 more punches than Burton (302-231) over the ten-round fight that legendary ringside commentator Teddy Atlas unofficially scored 97-92 (eight rounds to two) for Augustus. Surprisingly, Burton won the fight via a split decision, with one judge seeing the fight nine rounds to one in favor of Burton (99-90); another had it seven rounds to
three for Burton (97-92), and one judge, having some connection to reality, scored the fight 94-98 for Augustus. Atlas yelled after this decision, “This is a travesty here…. This is a disgrace…. This is what’s wrong with boxing. This is what chases our great fans away from this great sport.” It is difficult to fathom how that wide a disparity could exist in the scoring of a boxing match.

Experiments have been conducted to make the judges’ scores public by announcing them at various times during the fight itself, so that the fighters involved are at least aware of how the judges see the fight DURING the fight and have the ability to adjust their game plans accordingly. Those experiments have not been adopted throughout the professional boxing ranks, leaving boxers to guess how they are doing (or more clearly, how the judges THINK they are doing) during the fight itself. Knowing how one is doing in the eyes of judges may have a great effect on a corner’s tactics as the fight progresses, certainly including the risk/reward analysis of whether a fighter should continue and fight aggressively or throw in the towel. The present article will examine the current state of judging in boxing, including ethical concerns related to issues of justice for the boxers and explore suggestions on how to improve the current judging system.

**How Boxing is Scored**

Scoring a boxing match is a difficult task that requires a high degree of cognitive complexity. Professional boxing matches are comprised of three-minute rounds for men and two-minute rounds for women, with championship bouts lasting up to 12 rounds. During each round, judges are observing data on a second by second basis. Not only are the boxers exchanging blows that last in the tenths of a second, but they are also making subtle, nuanced maneuvers to gain offensive and defensive tactical advantages. While scrutinizing the action, judges are determining a winner of each round using the 10-Point Must system. A judge awards 10 points to the winner of each round and nine points to the loser; however, boxers can also lose points for getting knocked down or by fouling an opponent. Adding to the complexity, judges must utilize convoluted criteria to decide who won each round by considering four distinct elements of a boxing match: clean punching, effective aggression, defense, and ring generalship.

In the first dynamic, clean punching, a judge will evaluate the number of punches that a fighter lands in the scoring area (the forehead to the waist, not including the arms or the back) and compare it with the opponent. It is generally accepted
that harder, cleaner punches are weighted more than softer punches. To illustrate, if Boxer A lands 15 power punches in a round and Boxer B lands 20 jabs, the judges will generally award the round to Boxer A. The judges have the difficult task of determining when a punch is causing more damage, which is often left open for subjectivity. Secondly, effective aggression is a measure of efficiency, and boxers are rewarded for aggressive behaviors only when they are viewed as effective. If a boxer is moving forward, throwing punches, but not landing clean blows, a judge would not award them extra consideration for a round. Conversely, if two boxers are landing punches at a similar rate, but Boxer A is acting in a more aggressive manner, the judges will likely award the round to Boxer A. The third element, defense, is scored in a similar fashion. If Boxer A is vastly superior at defense than an opponent, Boxer A should receive more consideration in a round. That being said, judges will reward a fighter for defense only if it sets up a boxer’s offense. If a boxer is simply avoiding punches and not attempting to counterpunch, judges will often award the round to the boxer’s opponent.

Finally, ring generalship is the degree to which a boxer implements a strategy to control the action of a bout. For example, some boxers utilize a defensively oriented counterpunching style where they attempt to make their opponents miss punches so they are open to a counterpunch, while other boxers employ a volume punching style where they attempt to overwhelm and tire their opponents with the number of punches they throw. If a counterpuncher was boxing a volume puncher, whoever implemented their style more successfully in the bout would be awarded more consideration by the judges. Since there are a multitude of styles and strategies that can be utilized by boxers, judges must be knowledgeable about each approach and be able to identify when an approach is being successfully implemented.

At the end of each round, the boxing judges will individually consider all of the data they collected and determine a winner. This decision is made in a matter of seconds; and generally, the judges do not reveal their scores until the end of the fight. This is called closed scoring and is nearly universally used in boxing. In the rare instances when open scoring is utilized, judges reveal their scores in between rounds to inform the boxers, their teams, and the audience of the actual score in the bout. Most often, the scores are not announced following each round; but, instead, there are periodic updates, usually following the 4th and 8th rounds. Although open scoring has received a fair amount of discussion and has been implemented on occasion by different boxing commissions and sanctioning bodies, the boxing community has resisted universal implementation.
Judging the Judges: What’s Wrong with the System

Although the 10 Point-must system is universally used in professional boxing, it is also problematic with notable issues that need to be addressed. Firstly, the scoring system is inherently based on a subjective interpretation of the four scoring criteria by each judge. Although there is a general understanding of how the scoring criteria are operationally defined, there are still missing elements necessary to objectively score a boxing round. For example, it is unclear whether the four criteria should be weighted the same, or if clean punching should be prioritized. It is also unclear whether a boxer can win a round in which he or she lands fewer blows but has a superior aesthetic to the blows that he or she lands. In other words, do boxing judges have the ability to judge a boxing round on artistic principles? Secondly, the judges do not reveal their scores until the end of the bout, preventing the boxers from making adjustments during the boxing match. Knowing the score is a fundamental aspect of nearly every sport, and this knowledge dictates how athletes make competitive decisions. Without this information, participants are simply flying blind. Finally, the judges have an excessively sophisticated task to complete in a short amount of time, due to the degree of information occurring in each round and the complexity of the scoring system. Therefore, judges rely on what Tversky and Kahneman (1974) identified as cognitive heuristics or mental shortcuts that allow for individuals to make complex decisions quickly to score each round. Although these shortcuts allow for greater efficiency, they are also prone for cognitive biases. 

Cognitive biases emerge from quick decisions, and judging in boxing is no exception. For example, a judge may have a preconceived notion of who will win a bout they are scoring, and as the boxing match takes place, the judge may experience a confirmation bias. According to Plous (1993), a confirmation bias is where an individual favors new information that confirms a preexisting belief. If a judge believes that Boxer A will beat Boxer B based on previously viewed performances of said boxers, then the judge will recognize the clean punches, effective aggression, defense, and ring generalship of Boxer A more than Boxer B, even if Boxer B was objectively superior to his or her opponent in these categories. Similarly, a judge may prefer a particular boxing style, and therefore, score more favorable for boxers who exhibit those styles in a bout. To illustrate, if a judge prefers a counterpunching style more than a volume puncher, the volume puncher will have a more difficult time being viewed objectively in the areas of effective aggression and ring generalship by said judge, ultimately, putting him or her at an unfair disadvantage in the fight. In a similar way, a judge may be misled by a representativeness heuristic during a fight. Tversky and Kahneman (1974)
explained that individuals make decisions by selecting options that best represent their stereotyped idea of what they are trying to identify; and therefore, if a boxer is representing those traits most associated with winning a boxing match, a judge may award him or her the victory.

Suppose that a hometown boxer has a fan-friendly aggressive style, exudes confidence, and is the current champion in his or her weight division. He or she may receive louder applause from the audience, lift his or her hands in the air, and shake their heads in contempt when hit by a clean punch (all signs associated with winning a boxing match), and therefore, lead a judge to score him or her more favorably. However, these characteristics have little to do with the criteria for scoring a round in the 10-Point Must System and could lead a judge to incorrectly assess the dynamics of the boxing match. Furthermore, it is often believed that judges will give close rounds to the current champion, simply because they hold the title.

Questions regarding the reliability and validity of subjective scoring in sports is not unique to boxing. In the 2004 Olympic Games, Yang Tae-Young, a South Korean gymnast, was demoted from a gold medal to his bronze due to a 0.1 judging error on his parallel bars routine. The error was discovered shortly after the event, and unfortunately, the decision was not able to be reversed due to the requirement that all scoring disputes be made before the end of the competition. The repercussions of the decision extended beyond the gymnasium. The South Korean government has awarded Olympic medal winners a monetary incentive consistent with the type of medal won; and therefore, Yang has received a decreased monthly stipend since the incorrect decision. Although this loss of income is undoubtedly an issue of justice, the justice concerns in boxing should be prioritized as the risks in boxing are inherently greater. Not only are there justice concerns related to the financial impact of the scoring of a boxing match, there are concerns related to the health of the boxers. Therefore, resolving these fundamental issues is paramount.

Finding an Ethical Framework

Given the amount of money at stake at the highest levels of professional fighting, as well as the health of the fighters themselves, this paper argues that professional fighting has a moral obligation to bring their judging system into the 21st century. Ethical principles such as justice, autonomy, and nonmaleficence demand this revision, as the current scoring procedures are systematically problematic.
It is important to begin this section of the paper by pointing out that we do not wish to engage with the important and ongoing conversation in the literature regarding whether or not boxing is or can ever be a moral undertaking. Whole shelves have been written about this topic, and we do not intend to amend them here. To take just one example, Robert Simon’s landmark book *Fair Play: The Ethics of Sport*, says of boxing that it involves the use of physical force with the intention of causing harm to another as an act of violence, so henceforth “boxing surely is ethically problematic,” (2010, 209). These positions, while worthy of consideration, are set aside as we believe our question is a significant one regardless of whether or not the sport of boxing is itself ethical. We instead have turned to a practice within the sport itself, attempting to continue the type of reform that C. D. Herrera discusses in their “The Moral Controversy Over Boxing Reform.” Herrera reminds us that boxing has already undergone massive changes in its thousands of years of history, for boxing was originally “staged in sandpits, with no time limits or weight categories,” (2002, 165). While the sport has evolved with changing social attitudes, Herrera reminds us that as a sport its core has remained the same.

Instead of defending boxing itself, this paper sets the focal point of our analysis on the way in which many matches are decided. Taking autonomy as our first principle to consider, we recognize that some have argued that simply allowing boxers to compete fits with this principle. Still, for an action to truly be autonomous necessitates an atmosphere of informed consent. In the case of boxing, informed consent is often understood to apply at the signing of a contract agreeing to fight, including the setting of various rules for the fight, such as the size of the ring and the weight of the gloves. We assert, however, that every new round offers an opportunity for the fighters at hand to re-affirm their assent. For this assent to be an informed consent in line with the respect of each fighter’s autonomy, it only makes sense that the fighter has a sense of how they are doing.

One must remember that the fighter has the most at risk when a professional fight takes place. The fighters risk life and limb in pursuit of making a living for themselves and their families, as part of an industry designed for the entertainment of fans and the personal profit of promotors. This risk is compounded when one considers that most professional fighters emerge from poor urban environments with boxing serving as one of the few paths out they see for themselves and their families. From a Utilitarian perspective, one might argue that allowing fighters and fans the chance to know the judges’ scores during a fight might lead to less exiting matches or less tension concerning the events in the ring; yet, that concern pales in comparison to
the risk/reward that the fighters themselves face. While as outsiders, we might view the proceedings as entertainment, it’s a sport and those competing within that sport should know the parameters of the altercation. Imagine a football game in which we did not know how many points a touchdown was worth and could not know until after the game was over. Knowing the score of the football game does not make that game itself less exciting. Given how much both football players and boxers risk in terms of chronic traumatic encephalopathy (CTE), there is a moral obligation to provide them all the relevant information so that they determine how they should proceed. This assertion follows from both utilitarian and deontological perspectives, for Kant would appear to support the argument that purposely withholding information from relevant parties cannot be universalized and furthermore risks using the boxers as mere means toward the entertainment of others. These moral principles not only give us reasons to be disturbed by how judges scores are kept secret but also increase our concerns about how judging biases sometimes give fighters an unfair decision.

A moral right to have access to the information of how well a fighter is doing may be drawn from an alternative reading of John Rawls’ two principles of justice, especially as it is expressed in *Justice as Fairness: A Restatement* (2001). In his second principle, Rawls asserts that those in the Original Position should only accept social and economic inequalities that are to the benefit of all. I assert that a similar rule can apply to information, regardless of whether or not it slightly reduces the tension and resulting enjoyment of fans. During a bout, the system purposely withholds information from those who could be most directly impacted by possessing that knowledge, resulting in a situation that is certainly not to their benefit. A concert is no less enjoyable if one knows the set-list in advance, and the ninth round of a boxing match can still be dramatic even if we know one fighter is significantly ahead on the scorecard.

**A Path Toward A More Just Judging System**

Due to the highlighted justice concerns inherent in boxing scoring, it is recommended that all levels of boxing commissions implement one primary and two secondary measures to improve the 10-Point Must System. First, open scoring is recommended to enhance transparency for the boxers during the bout. Boxers should have the right to know how the judges are scoring a fight, so they can make informed decisions about when to box more aggressively and risk their health for a greater chance of winning and when to box more defensively and protect their physical health. Boxers also should be made aware of instances when a boxing strategy is not being
rewarded by the judges, so they can adjust their approach to obtain a more favorable outcome.

To further illustrate the benefits of open scoring, readers should consider the 1999 boxing match between Oscar de la Hoya and Felix Trinidad. In a highly anticipated bout between two undefeated fighters, Felix Trinidad defeated Oscar de la Hoya after de la Hoya switched from an aggressive strategy to a defensive approach late in the fight. Thinking that he had won enough rounds early in the fight, de la Hoya elected to forfeit the final few rounds by focusing solely on defense in order to ensure that Trinidad would not knock him out late in the bout. What de la Hoya did not know was that the judges’ scores were very close, and that he did not have a sufficient lead to give away the late rounds. Trinidad ended up winning the fight and keeping his undefeated record. It is impossible to overlook the possibility that if de la Hoya knew he was in a close fight, he could have continued to fight aggressively and potentially defeat Trinidad.

Another primary example supporting open scoring can be found in the 2018 heavyweight championship fight between Deontay Wilder and Tyson Fury. Going into the final round, Wilder needed to knockdown Fury and secure a 10-8 round to obtain a draw and retain his title. Fury, not knowing the score, left himself open for a devastating punch from Wilder that put him on the canvas. The extra point from the knockdown led to a 10-8 round for Wilder and produced a draw. If Fury had been aware that losing the final round 10-9 would have guaranteed him the victory, he could have been extremely cautious and not only won the bout but also avoided the extra damage from Wilder’s vicious punch. Clearly, knowing the score would have had an impact on a fight that ended in another controversial decision, with one judge scoring it 115-111 for Wilder, and another judge scoring it 114-110 for Fury, and the decisive judge scoring it a draw at 113-113.

Although there are many examples supporting open scoring, the general boxing public has been opposed. In a 2018 article, Sares demonstrated a lack of support to change the current system to open scoring in his survey of boxing insiders. The arguments presented predominately supported the experience of the boxing public, as it was postulated that open scoring decreases drama and promotes cautious behaviors when a boxer is decisively winning a bout. At the core of these arguments is the concern that open scoring will change the essence of boxing by making it less exciting. That being stated, one could argue that requiring boxers to wear padded gloves decreased the excitement of the sport; however, few would suggest that bare knuckle boxing has a place in modern sport. Open scoring should become a similar
phenomenon simply due to our current understand of brain trauma. Under open scoring, boxers will be able to decide when to protect themselves from unnecessary punishment. Altering a sport for the safety of the participants should never be viewed as unacceptable.

There were also concerns noted by Sares (2018) that open scoring puts boxing judges in a dangerous position if an audience disagrees with their decision. However, the article failed to recognize that the boxers are putting themselves inherently at risk, and therefore, should have the right to be able to decide if they want to know the score throughout the fight. One suggestion to mitigate the risk to the judges would be to provide scoring updates only to the boxers and their trainers. By excluding the audience from knowing the score, not only would the safety of the judges be intact, but also the dramatic announcement of the final scores would remain. Since the justice issues noted in the present article are directly related to the boxers, it is irrelevant if the audience is aware of the scoring prior to the end of the fight.

A secondary recommendation is that the definitions of the four boxing scoring criteria (clean punching, effective aggressiveness, defense, and ring generalship) be revised and enhanced to produce more objectivity among the judges. Indeed, simply by clarifying if the four criteria are to be scored equally or if one or more criteria should receive stronger consideration by the judges, the scoring would be more reliable and valid. With this in mind, further improvements could be made by exploring the interpretation of the scoring criteria guidelines among the population of licensed boxing judges. This exploration could uncover additional discrepancies that could be addressed through further clarification of the scoring definitions.

A third recommendation, which could be used in accordance with open scoring, suggests that boxing increase the number of judges involved in scoring a fight. Instead of having three judges on three of the ring’s four sides, boxing could instead have five judges, with one on each side of the ring and a fifth watching a closed-circuit camera angle from directly above the ring. This would help prevent times when the angle of an altercation in the ring prevents the majority of judges from seeing the action. In addition to having five judges, the final decision could involve having the two most extreme scores (high/low) excluded from the final decision.

Finally, it is recommended that boxing judges take steps to manage the cognitive biases that may impact the objectivity of their scoring similarly to the way qualitative researchers manage their biases. According to Creswell and Poth (2018), qualitative researchers should use active trustworthiness strategies to address their implicit biases during an inquiry. One pertinent strategy that could be used by boxing judges is to
Gray and Russ

write an epochè or an account of their personal beliefs and experience that could impact a particular judging assignment directly prior to judging the bout. By exploring and identifying ones’ implicit biases, they will increase awareness, and hopefully, bracket or suspend the preconceived presuppositions of how a boxing match will be determined. It would be more productive to identify and address potential biases as opposed to simply striving for objectivity without an established plan.

Conclusion

We conclude that these revisions to the process of judging boxing matches allow for the event to transpire in a more transparent and ethical manner. Imagine if a baseball umpire did not announce whether a pitch was a ball or a strike until after a batter strikes out or walks. Not only would batters be unable to adjust their approach during the at-bat, but the pitchers would have no idea on where to throw their pitches. Although this idea seems ridiculous, not knowing the score during a boxing match is considered normal. Boxers are unable to adjust their strategies or tactics throughout a bout, and the audience is often left puzzled as to why the judges’ verdict is different from their own perception of who won a bout. Through scoring reform, boxing will begin its transition into the 21st century. Without changing the nature of the sport itself, boxing could instantly be made more ethical and just.

References


Sares, Ted. 2018. “Are you in Favor or against open scoring in boxing? Results of a TSS Survey.” *The Sweet Science*.

Difficult Circumstances: Situationism and Ability

Marcela Herdova and Stephen Kearns
Florida State University

Biographies
Marcela Herdova is Assistant Professor of Philosophy at Florida State University. She previously worked as Research Associate on the “Self-Control and the Person: A Multi-Disciplinary Account” project at King’s College London and as Postdoctoral Research Fellow in Free Will and Self-Control at Florida State University. Her research interests include action theory, free will, moral psychology, consciousness and applied ethics. Stephen Kearns is Associate Professor of Philosophy at Florida State University. His research interests include topics in action theory, metaphysics, epistemology, moral philosophy, the philosophy of mind, and others.

Publication Details

Citation
Difficult Circumstances: Situationism and Ability

Marcela Herdova and Stephen Kearns

Abstract
Certain aspects of our situations often influence us in significant and negative ways, without our knowledge (call this claim "situationism"). One possible explanation of their influence is that they affect our abilities. In this paper, we address two main questions. Do these situational factors rid us of our abilities to act on our sufficient reasons? Do situational factors make it more difficult for us to exercise our abilities to act for sufficient reasons? We argue for the answer 'sometimes' to both these questions. We then explore the consequences of this view for moral responsibility.

Keywords
Situationism, Circumstances, Abilities, Moral Responsibility

1. Introduction
It is now widely recognized that the presence of different situational factors can affect our behavior in surprising, and sometime negative ways. Often, we are not aware that our behavior is so influenced. What’s more, situational factors cause us to act in certain ways without providing us with reasons to act in such ways (or, at most, they provide only weak reasons). These claims capture what we mean by the thesis of situationism (other definitions of this term may have stronger claims in mind than ours). The relationship between situationism and various important philosophical topics has been explored by many thinkers. For example, Doris (2002) and Miller (2013, 2014) argue that the situationist data cast doubt on the idea that humans have virtues or vices. Further, Nelkin (2005), Nahmias (2007) and Vargas (2013) explore the extent to which the situational factors affect, respectively, our freedom, autonomy and moral responsibility. In Herdova and Kearns (2015) we link situationism with an interesting type of moral luck, and in Herdova and Kearns (2017) we examine how our situations may affect the extent to which we are reasons-responsive.

In this paper, we explore how situationism affects agents’ abilities. In particular, we shall address the following two questions. Do situational factors (such as those studied in situationist experiments) rid us of our abilities to act on our sufficient reasons? Do situational factors make it more difficult for us to exercise our abilities to act for sufficient reasons?
We structure the paper as follows. In section 2, we present some relevant data concerning the impact of situational factors on our acting for sufficient reasons. In essence, there is good evidence that the majority of people fail to act on their sufficient reasons when exposed to certain situational factors. In section 3, we introduce two main ways in which appealing to agents’ abilities may explain the situationist data: situational cues may eliminate our abilities or make them harder to exercise. In section 4, we argue that situational factors do not render agents generally unable to act on their sufficient reasons, but that certain agents may not be free to act on their sufficient reasons. In section 5 we argue that situational factors sometimes make it more difficult for us to exercise our abilities to act on our sufficient reasons. After considering the consequences of this for moral responsibility in section 6, we end on a somewhat brighter note in section 7, with the suggestion that some situational factors make it easier to exercise our abilities to act on sufficient reasons.

2. Situationist Data

There are many different types of situationist experiments. Below we introduce the results of some such significant experiments which feature sufficient reasons (seeing as it is the abilities to act on those reasons with which we are concerned). By “sufficient reasons” we understand those reasons which entail obligations: if S has a sufficient reason to A, S ought to A.

One notable group of experiments documents the so-called bystander effect, according to which the likelihood of one’s intervening in an emergency situation depends, in part, on how many other people witness this emergency situation. Specifically, the more people present in an emergency situation, the less likely it is that any of those people will intervene. For instance, Darley and Latané (1968) conducted an experiment where the participants witnessed an (apparent) medical emergency (they overheard a staged epileptic attack). Out of those who believed that the attack was overheard also by four other people, only 31% offered assistance in the relevant timeframe (before the person having the apparent attack was cut off after 125 seconds). However, the number of subjects who intervened was significantly higher in a condition where the subject believed he was the only one to overhear the attack; in this experimental condition, 85% of the subjects intervened.

1. In section 3, we explore why such abilities are particularly interesting from a philosophical perspective.
Latané and Darley (1968) also tested for the bystander effect in a non-medical emergency situation (smoke filling up a room). Out of those who witnessed the smoke on their own, 18 out of 24 subjects reported the smoke. In trials where the experimental subjects were accompanied by two passive confederates, only one in 10 subjects reported the smoke. In both of the above experiments, whether the subjects helped or otherwise intervened largely depended on their being accompanied or not.

Whether one provides help or intervenes in pressing situations seems to be influenced by situational factors other than the number of people present. For instance, in the Good Samaritan experiment conducted by Darley and Batson (1973), seminary students were asked to give a talk in a nearby building. On their way to give the lecture, they came across a person in apparent need of medical attention. Some students were told they were running late; only 10% of the students in this group offered assistance. On the other hand, out of those subjects in a low-hurry condition, who were told they had enough time, 63% of the subjects helped. The students did also differ, aside from how much time they had, in the content of their lecture: some were going to talk on the parable of the Good Samaritan and some on job prospects. However, unlike the hurry factor, the lecture content did not to make a significant difference with regards to whether they offered assistance or not.

One final set of experiments we will mention here is Milgram’s obedience experiments (1963, 1974). These focused on studying subjects’ behavior under the influence of authority. Subjects, however, believed that they were taking part in a learning experiment and were asked to deliver apparent electric shocks to “learners” upon them providing wrong or no answers to the relevant questions. The experimental subjects used a range of 30 levers for this purpose, each associated with a different degree of shock (the highest apparently being 450 volts). In Experiment 1, approximately two-thirds of the subjects complied with the instructions of the experimental confederate, and continued to deliver shocks all the way (that is, pulling all the levers, including the one delivering the highest degree of shock). The subjects continued to increase the voltage despite the fact that after the 20th question, the confederate apparently receiving the shocks would bang on the wall and then stop providing answers.

In Milgram’s Experiment 11, subjects were free to choose what levels of shock to deliver (the confederate emphasized that they could use any such levels). The vast majority of the subjects delivered the lowest shocks when the choice of shock severity was left up to them. This indicates that the presence of the experimental confederate acting as an authority figure played a significant part in whether subjects delivered high
degrees of shocks to the learners. The obedience experiments were partially replicated by Burger (2009) who found that “obedience rates in the 2006 replication were only slightly lower than those” in Milgram’s original experiments (2009, 1).

While the above list only provides us with a relatively small sample of the situationist experiments, it gives us a good indication of the impact that some (normatively irrelevant) situational cues may have on our behavior. What we may do (or refrain from doing) in different scenarios largely depends on the presence of arbitrary situational factors (such as being in a hurry, being accompanied by people, or being asked by authority figures to act in certain ways, and many others).

3. Situationism and the Ability to Respond to Sufficient Reasons

The situationist data suggest that, due to the influence of situational factors, people very often fail to act on their sufficient reasons (i.e., those reasons which entail obligations to perform actions). No one should pull all the levers in the obedience experiments, and yet many do (because they are firmly asked to do so). Everyone who witnesses the apparent emergencies in the bystander experiments should alert someone, yet many do not (because they are accompanied). Every seminarian should help the stranger in the Good Samaritan experiment, yet almost all of those in a hurry do not. These results very plausibly generalize to people outside experimental scenarios, given the strength of the data, the fact that the experimental subjects were assigned their groups on a random basis, and the fact that these subjects were not chosen for the experiments on the basis of their susceptibility to situational factors. That is, the data above (and myriad other data from the situationist literature) strongly suggest that all of us are very often significantly affected by the presence of various situational factors. These factors frequently prevent us from acting in ways we ought to act.2

But how might the situational factors often prevent us from acting on our sufficient reasons? Most obviously, we may wonder whether situational factors such as those above affect agents’ abilities. One explanation of why people don’t respond to sufficient reasons is that the presence of certain situational factors makes it more challenging for them to do so (resulting in, ultimately, many people failing to do so). Even more worrying is the possibility that due to situational factors, people can’t act on

2. Some may claim that the subjects of the above experiments do not have sufficient reasons to help/alert authorities/etc., because no one is really in medical need, there is no fire, etc. This claim is highly questionable (shouldn’t we act in the face of possible medical need, potential danger, etc.?), but even if it is right, the experiments show us how people would act when faced with actual sufficient reasons.
their sufficient reasons. That is, perhaps situational factors take away our ability to act on our sufficient reasons.³ This is in part what concerns Dana Nelkin (2005):

many of [the experimental results] seem problematic because the subjects in them don’t seem to be acting for good reasons, or at least their behavior raises a question about whether they are. And further, the way in which the subjects seem to proceed raises a question about whether they can act for good reasons—in some important sense of “can.” (Nelkin 2005, 200-201)

Manuel Vargas suggests (2013) that situationism shows that we have various capacities (to act on reasons, to assess reasons, etc.) only in certain situations—situational factors can rid us of such capacities (and thus related abilities). Vargas contemplates the suggestion that, while this may be so, our basic capacities (to respond to reasons, etc.) remain constant across situations. To this he says:

Even if our basic capacities are stable across contexts, our abilities to exercise them vary by circumstance... (Vargas 2013, 334)

This certainly seems to be an explicit commitment to the idea that situational factors may rid us of abilities (Vargas later concludes that it is at least a significant possibility that some subjects in certain experiments, such as the Good Samaritan Experiment, “suffered a loss of free will” (2013, 339) precisely because the relevant situational factors may undermine subjects’ reason-related abilities).

Indeed, the idea that this is what situational factors do is worth exploring in detail. One reason to take these hypotheses seriously is that they are relevant to questions of moral responsibility: if situationism shows that some of our abilities are strongly affected by situational factors, perhaps we ought to reassess our moral practices of praise and blame. Susan Wolf thinks of both freedom and responsibility as requiring the ability to act on the True and the Good. Wolf takes this to entail that “an agent is responsible if and only if the agent can do the right thing for the right reasons” (Wolf 1990, 68). If situational factors strip us of our ability to act on sufficient reasons (which we take to be at least roughly equivalent to “right reasons”), then, on Wolf’s

³. Of course, the data do not suggest that agents influenced by situational factors are unable to do otherwise than any action they perform (or that every ability of the agents’ is harder to exercise). For all the data show, a subject in the Milgram experiment is free to pull the lever with his left hand, or instead with his right; a hurrying seminarian may be able to run to his destination or merely walk quickly. What is of interest is whether situational factors adversely affect agents’ abilities to act on sufficient reasons.
picture, they also strip us of our freedom and our responsibility. A plausible extension of Wolf’s view is that, the more difficult it is for an agent to exercise her ability to act on her sufficient reasons, the less morally responsible she is for failing to do so. In what follows we argue that situational factors do indeed sometimes rid agents of the ability to act on sufficient reasons (section 4), and that situational factors sometimes make exercising these abilities more difficult (section 5). We examine the implications of these claims for moral responsibility in section 6.

4. Do Situational Factors Eliminate Abilities?

Let us start with the worry that we might lack the relevant abilities altogether. Take, for instance, the ability to respond to a sufficient reason to help. One extreme version of the worry is that situationism shows that *everyone* lacks the ability to respond to sufficient reasons to help in what may be a wide range of circumstances—those which involve certain salient environmental factors. However, the data do not support this concern. The response of any select group of people in the aforementioned experiments, and indeed in situationist experiments across the board, is far from uniform. Consider, again, the Good Samaritan experiment. The data for this experiment do not show that *all* agents lack the ability to respond to their sufficient reason to help. Recall that even in the high-hurry condition, there are still agents (10% of the group) who do offer to help. So, it is not the case that the relevant situational factors in this scenario erase everyone’s ability to respond to such sufficient reasons.

A more plausible worry one might have is that the data support the claim that a *significant number* of people, in certain situations, lack the ability to respond to sufficient reasons, due to the influence of situational factors. If this is true of *experimental subjects*, it is most likely true of *us*—many of us lack the ability to respond to sufficient reasons in certain situations.

4.1 Situational Factors and Accounts of Abilities

Though the above hypothesis, if true, would explain why many people don’t respond to sufficient reasons, it is not obviously the best explanation available. The situationist data, in essence, highlight that people’s behavior in the face of similar situational factors often conforms to certain patterns—the data do not on their own imply that these patterns track the limits of our abilities. Indeed, there is some obvious evidence that subjects who fail to act on their sufficient reasons due to situational factors can act on such reasons. For example, subjects who do not help are not
thwarted in their attempts to help—they don’t even try to help. They are not physically incapacitated by their situations, nor overcome with irresistible desires. The evidence concerning what they ought to do is not cleverly hidden from them—for example, the seminarians have a clear view of the person sitting at the side of the road; the Milgram subjects receive sufficient auditory and/or visual feedback indicating the learner’s distress. Though we cannot conclusively rule out that many subjects are unable to respond to sufficient reasons in specific situations, the evidence for this (from the situationist data) initially seems lacking.4

We should not dismiss the claim that situational factors rid us of our abilities to act on our sufficient reasons simply on the basis of the above considerations. After all, the thesis that situational factors render many of us unable to act on our sufficient reasons can be adequately assessed only once we are clear what we mean by “able”. Few would deny that we retain (in the face of adverse factors) the general ability to act on the sufficient reasons we have. Just as a person locked in a small room may still have the skill, know-how, physical capacity, and willingness to drive a car, and thus in some sense is able to drive a car (i.e., has the general ability to drive a car), there is a clear sense in which she currently cannot. Given our interest in moral responsibility, and whether situationism undermines it, the most obvious way to interpret the claim that situational factors eliminate the ability to act on sufficient reasons is that they eliminate the freedom to act on sufficient reasons. Certainly the person locked in her room is not free to drive a car.

Do situational factors rid a significant number of us of the freedom to act on sufficient reasons? The considerations we mention above (e.g. these factors do not physically incapacitate us, nor produce irresistible desires in us, etc.) is not conclusive evidence that they do not rid us of such freedom. Perhaps—some might say—these factors leave us with our general ability to act on sufficient reason, but eliminate our freedom-level ability to do so. In what follows, we examine various understandings of freedom-level ability, with the aim of understanding whether situational factors may strip us of the freedom-level ability to act on sufficient reasons.

One popular understanding of the freedom to do otherwise is the so-called “conditional analysis” (see, for example, Ayer (1954), Bok (1998)). Though the details differ, this analysis states:

4. Indeed, that some people do respond to sufficient reasons in situations where most don’t is some evidence that even the people who do not respond to the relevant reasons could have done so.
S has the (freedom-level) ability to A if and only if S would A if S tried/chose/intended to A.

Though this analysis has become less popular since its heyday, similar analyses in terms of dispositions (rather than counterfactuals) have recently been propounded. Thus Michael Fara (2008) claims:

S has the ability to A in circumstances C iff she has the disposition to A when, in circumstances C, she tries to A. (Fara 2008, 848)

Kadri Vihvelin proposes a similar account of ability, but also believes that a counterfactual analysis of dispositions, along the lines of the one defended by David Lewis (1997), may be given. Combining her dispositional account of ability with a Lewisian account of dispositions, Vihvelin (2004) arrives at the following analysis of ability:

Revised Conditional Analysis of Ability: S has the ability at time t to do X iff, for some intrinsic property or set of properties B that S has at t, for some time t’ after t, if S chose (decided, intended, or tried) at t to do X, and S were to retain B until t’, S’s choosing (deciding, intending, or trying) to do X and S’s having of B would jointly be an S-complete cause of S’s doing X. (Vihvelin 2004, 438)

The details of Vihvelin’s account (or indeed Fara’s) need not concern us here. The pertinent point is that such accounts share an important similarity—they understand the ability to do something as related to a (hypothetical) attempt to do it. Very roughly-speaking, on each of these accounts, an agent is free to perform an action if (but not only if) it is true that, if the agent were to try to perform this action, the agent would successfully manifest her disposition to perform it.

At first blush, it seems clear on these accounts of (freedom-level) ability that many of those situationist agents who do not act on their sufficient reasons to help someone could have so acted. For instance, it seems plausible that (many) subjects in Milgram’s experiments would have refrained from (seemingly) shocking someone had they tried. Many or most (if not all) seminary students in a hurry would have helped the prone figure had they attempted to do so. After all, the situations that the subjects find themselves in do not feature obstacles which would likely make such attempts...
unsuccessful—acting in these ways in such situations does not require unsurmountable amount of either physical or psychological strength.\textsuperscript{5}

May we conclude, then, that on such accounts of freedom-level ability, situational factors do not remove agents’ abilities to act on their sufficient reasons? There is reason to be cautious about concluding this. Recall that our main question is this: are agents who are adversely affected by situational factors free to act on their sufficient reasons? Given the above analyses of freedom to do something, this question becomes (roughly): would agents who are adversely affected by situational factors manifest a disposition to act on their sufficient reasons were they to try to act on their sufficient reasons? Even though it may be true that subjects would, for example, help someone were they to try to do so, it does not follow that they would act on their sufficient reasons were they to try act on their sufficient reasons (even if what they had sufficient reason to do is to help someone). To illustrate this point, consider the following possibility. Situational factors (sometimes) render us unaware of our sufficient reasons (and of the fact that we are unaware of them). Thus were we to try to act on our sufficient reasons, we would (sometimes) fail to do so (because we would instead unknowingly act on the basis of things which are not sufficient reasons). This is so even though, were we to try to help the person in need, we would succeed in so doing (because, were we to try to help the person in need, we would perforce be aware of his existence!). In essence, then, given that we accept the above accounts of freedom-level ability, if situational factors prevent us knowing our sufficient reasons, then they also render us unfree to act on them (whereas, if situational factors do not prevent our knowing our sufficient reasons, they do not render us unfree to act on them).

The above dispositional/conditional account at best leaves it open, then, whether situational factors remove our freedom to act on our sufficient reasons (depending on whether such factors render us ignorant of our sufficient reasons). But we may, of

\textsuperscript{5} Might someone claim that the subjects adversely affected by situational factors lack the \textit{disposition} to, for example, help someone in need? Thus, though the subjects would help were they to try, this is only because the closest possible world in which they try is also one in which they have the disposition to help. If this is right, then on dispositional accounts such as Fara’s and Vihvelin’s, they are not free to act on their sufficient reasons to help. But we have already argued that the subjects have the \textit{general} ability to help. This general ability is, in essence, a disposition to help on the basis of sufficient reasons. On dispositional accounts of (freedom-level) ability, an agent is \textit{free} to help if they are thus disposed and are placed in the right circumstances. Roughly-speaking, a circumstance is right if, were the agent to try to help, they would manifest their disposition to help. Such circumstances obtain, we have argued, even when agents are adversely influenced by situational factors.
course, reject such accounts. One obvious reason we might worry about these accounts is that situational factors may render agents unable to even try otherwise than they do, in which case they plausibly cannot do other than they do either (an implication such accounts seem to deny). Another worry is that the above accounts are clearly compatibilist-friendly—that is, these accounts allow that determinism does not rule out the ability to act otherwise.6 This is, to say the least, a controversial assumption. Let us now consider, then, an incompatibilist account of ability and its relation to the situational data.

One who rejects compatibilism (i.e., an incompatibilist) is more likely to accept something like the following account of freedom-level ability:

\[
S \text{ has the (freedom-level) ability at } t \text{ to } A \text{ if and only if it is compatible with the laws of nature and past up until } t \text{ that } S \text{ As.}
\]

Do the situationist data cast doubt on our ability to act on sufficient reasons conceived in this way? Not as far as we can see. It is entirely consistent with the data that the choices the subjects make, the best judgments they make, and the considerations that occur to them, are all undetermined. (In other words, it is compatible with the laws of nature and the past up until the relevant time (e.g. the time of choice, judgment, etc.) that subjects make different choices, best judgments, etc.) Thus even those subjects who fail to act on their sufficient reasons might act on those sufficient reasons in a possible world with the same past and laws of nature as this world, which is (according to the above incompatibilist account) sufficient for them having the freedom-level ability to act on a sufficient reason.

This is easy enough to see. Take, for example, the Darley and Latané (1968) bystander experiment, in which subjects overheard an apparent epileptic-like fit. Someone who believes the subjects’ choices were undetermined in these experiments may hypothesize that there was around an 85% objective probability of the unaccompanied subjects choosing to help, while the probability of the subjects who believed that the attack was overheard by four other people choosing to help was around 31%. This would nicely explain why 85% of the former chose to help, and only 31% of the latter so chose. Of course, it is far too simplistic to judge the objective probabilities of a subject helping from the percentage of those in his group that helped. Still, the point is that the hypothesis is perfectly consistent with the data (indeed, not simply consistent with it, but one possible explanation of it).

---

6. Determinism is the thesis that the past and the laws of nature entail every truth.
We should not, however, conclude from the above that the incompatibilist can unproblematically accept that situationist factors pose no threat to our freedom. First, such considerations simply show that the situationist data are *in harmony* with the presence of indeterminism—it does not (and cannot) show that situational factors do not in fact determine agents’ behavior. Second, and more importantly, there is good reason for incompatibilists to reject the simple account of freedom-level ability given above. Even if the performance of a certain action is compatible with the past and the laws, it does not follow that the agent is free to perform this action. If, at t, an agent has no idea how to A and is unable to voluntarily or intentionally discover how to A, she is not free at t to A after t. It may still be compatible with the past and the laws up until time t that she As after t, however, because it may be compatible with the past and laws that she *accidentally* learns, after t, how to A. This suggests that the incompatibilist account of freedom-level ability must be modified roughly as follows:

\[ S \text{ has the (freedom-level) ability at } t \text{ to A if and only if it is compatible with the laws of nature and past up until } t \text{ that } S \text{ As and } S \text{ knows at } t \text{ how to A (or is able intentionally to discover how to A).} \]

Just as with the dispositional/conditional account of freedom-level ability, epistemic considerations prove relevant to determining whether situational factors remove our freedom to act on sufficient reasons *as this freedom is thought of by the incompatibilist*. If situational factors render us unaware of our sufficient reasons (or that we have them), they plausibly also prevent us from knowing how to act on them, and thus render us unfree to act on them.

Indeed, these kinds of considerations suggest an argument that is independent of whatever account of freedom-level ability we accept that situational factors can and do prevent our being free to act on our sufficient reasons. To this argument we now turn.

### 4.2 Situationism, Abilities, and Ignorance

In this section we present what we take to be the best argument for the claim that situational factors remove our freedom to act on sufficient reasons. Roughly, the idea is that situational factors (sometimes) render us *unaware* of our sufficient reasons. If an agent is unaware of her sufficient reasons, she cannot act on them. So, though she may have the *general* ability to act on her sufficient reasons, she is not *free* to do so (i.e., she lacks the freedom-level ability) due to lacking certain relevant knowledge (of her sufficient reasons). What should we make of this argument? This initial version has two premises. First, that situational factors prevent our being aware of sufficient
reasons. Second, when we are so unaware, we are unable to act on such reasons. We shall explore each premise in turn (and, in so doing, formulate a more sophisticated version of the argument [in 4.2.2]).

4.2.1 Awareness and Unawareness of Sufficient Reasons

Are the subjects in the discussed situationist experiments aware of their sufficient reasons? For obvious reasons, we shall focus on those subjects who fail to act on their sufficient reasons due to the influence of situational factors. These include subjects who do not offer medical assistance when it is needed (due to their being in a hurry), fail to report an emergency (due to their being accompanied by others), and deliver apparently lethal electric shocks (due to their being asked to do so by an authority figure). Do these subjects fail to recognize their sufficient reasons, or do they recognize but fail to react to these reasons? Two extreme hypotheses are available. One is that the situationist subjects who, due to the relevant factors, fail to act on their sufficient reasons, do so because they always fail to recognize their sufficient reasons. The other is that such agents always recognize their sufficient reasons, yet fail to act on them regardless. Below we suggest that the available data, based on agents’ subjective reports, do not support either of the extreme hypotheses. In essence, sometimes situational factors affect our recognition of reasons, and sometimes they affect our reacting to them.

First, it should be noted that whether or not experimental subjects recognize their sufficient reasons is difficult to settle due to methodological issues. Not all situationist experiments have relevant post-experiment interviews, and those that do are often not detailed enough. This is particularly important (and problematic) given that a failure to recognize a reason might occur in different ways. It may be the case that the subjects fail to recognize the relevant reason as a fact. For example, subjects might fail to recognize a slumped person who appears to be in distress as needing help. Further, subjects may fail to recognize that the relevant fact is a reason to act; for example, they might fail to recognize that the fact that a man is in need of medical attention is a reason for the subject to help. Lastly, it may be that while the subjects do recognize a certain fact as a reason to act, they do not recognize this as a sufficient reason to act. So, while the subject might recognize that the fact that there is someone who appears to be in need of medical attention is indeed a reason to help, the subject might not recognize that it is a sufficient reason—as something she ought to act on.

Another problem concerns the reliability of self-reports. We might expect some unintentional confabulation (given that subjects are asked to report on their attitudes
after—possibly highly charged—events). Intentional confabulation is a problem too since subjects may attempt to rationalise or hide what could be viewed as a socially unacceptable behavior (neglecting to offer help, delivering apparently painful shocks etc.).

While keeping in mind the problems concerning reliability, we can extrapolate a number of hypotheses about subjects’ recognition of their sufficient reasons from the data. In the bystander experiment that concerned reporting the presence of smoke, some subjects seemingly failed to interpret the relevant reason as even a fact. When the subjects were asked by the interviewer if they encountered any difficulties while in the waiting room, most subjects mentioned the smoke. They were then further prompted to explain what happened. Latané and Darley state that:

Subjects who had not reported the smoke … uniformly [italics added] said that they had rejected the idea that it was a fire. Instead, they hit upon an astonishing variety of alternative explanations, all sharing the common characteristic of interpreting the smoke as a non-dangerous event. (Latané and Darley 1968, 219)

If this is indeed how the subjects who failed to report the smoke viewed the situation, then this suggests that they failed to be appropriately receptive to reasons—they failed to recognize that there was a potentially dangerous event occurring which needed to be reported. What is striking about this particular case is that, according to the experimenters, all of the subjects who failed to report the smoke interpreted the situation in a similar fashion: as something not dangerous. (Still, as noted previously, we need to be cautious regarding the accuracy of such subjective reports—some subjects were, perhaps, trying to save face.)

Not all the experiments are like the one above, however. There are other experiments where it is relatively clear that subjects who failed to act on their sufficient reasons were indeed aware of such reasons. Take the obedience experiments, for example. It is already very plausible that many of the subjects recognized that a) the shocks apparently caused extreme pain to someone, b) this fact is a reason to stop causing pain to another person, and c) this reason is sufficient, but these subjects decided regardless to act in line with the requests of the confederate urging them to carry on with the experiment. Some evidence for this comes from observations about how the subjects behaved during and after the experiment. Milgram explains that the procedure created “extreme levels” of nervous tension:
Many subjects showed signs of nervousness in the experimental situation, and especially upon administering the more powerful shocks. In a large number of cases the degree of tension reached extremes that are rarely seen in sociopsychological laboratory studies. Subjects were observed to sweat, tremble, stutter, bite their lips, groan, and dig their fingernails into their flesh. These were characteristic rather than exceptional responses to the experiment. … Fourteen of the 40 subjects showed definite signs of nervous laughter and smiling. … Full-blown, uncontrollable seizures were observed for 3 subjects. (Milgram 1963, 376)

After the experiment, when the maximum shocks had been delivered:

many obedient subjects heaved sighs of relief, mopped their brows, rubbed their fingers over their eyes, or nervously fumbled cigarettes. Some shook their heads, apparently in regret. (Milgram 1963, 376)

This level of distress can plausibly be explained by the conflicting reasons that the subjects had. Even if they did not recognize the relevant fact as a sufficient reason to stop delivering the shocks, their distress levels indicate that they, minimally, recognized it as a reason—as something deserving further consideration. Presumably, if the subjects did not recognize the learner’s (apparent) distress as a reason for (altering their course of) action, creating a conflict with other reasons they had (stemming from the requests of the confederate), they would not have displayed such tense behavior. In other words, if they had viewed the learner’s reactions as morally irrelevant, and as something that did not clash with other reasons they had, they would not have reacted in such a strong way.

In sum, then, it seems to be the case that sometimes agents fail to recognize their sufficient reasons due to situational factors, and sometimes they fail to act on their reasons due to situational factors. Both extreme hypotheses mentioned above are false—situational factors may affect us epistemically or non-epistemically.7

How does this affect the argument that situational factors strip agents’ abilities to act on sufficient reasons because they render agents unaware of these sufficient reasons? In effect, the above considerations show that this argument is somewhat limited in scope. Not all situational factors render subjects unaware of their sufficient reasons.

---

7. See Herdova 2016 for further discussion of the evidence that subjects in situationist experiments are often unaware of their sufficient reasons.
reasons, and even those that (arguably) do, often affect only a subset of subjects in that way. Still, we may ask the question whether those subjects who are unaware of their sufficient reasons are able to act on these reasons. To this question we now turn.

4.2.2 Unawareness and Ability

Let us assume, then, that some agents are rendered unaware of their sufficient reasons due to the influence of situational factors. Are at least these agents unable to act on their sufficient reasons? To answer this question, we need to answer another, namely: is it true that one cannot act on a sufficient reason if one is unaware of it? There is perhaps a reading of this claim on which it is true but relatively uninteresting and irrelevant (for our present purposes). That is, it is no doubt true that one cannot act on a sufficient reason while at the same time remaining unaware of it (in acting on a sufficient reason, one is also thereby aware of the sufficient reason). An agent may lack this ability, however, but still have the ability to act on a sufficient reason (even if the agent is ignorant of her sufficient reasons). Consider the following example. In some sense it is true that one cannot drive a car if one is not in a car. We can spell out this sense as follows: one cannot drive a car while at the same time not being in car. Still, even if one is not in a car, one might still be able to drive a car. This is straightforwardly true of having the general ability to drive a car (one might have the requisite skills, know-how, etc., without being in a car), but it is also true of the freedom-level ability to drive a car. Even if one is not currently in a car, one may still be free to drive a car given that one has access to a car, the ability to get into it, the ability to drive it, access to a road, etc.

Similarly, then, one may be free to act on one’s sufficient reasons even if one is unaware of them, given that one has ready epistemic access to these reasons, the (freedom-level) ability to seek out, discover, and reflect on these reasons, the ability to act on them once they are known, the opportunity to perform the relevant action, etc. Consider an agent who does not know her sufficient reasons but is perfectly free to figure out what they are (and knows she is so free). This agent is also free to act on her sufficient reasons. Of course, she cannot do so right away, given her epistemic position, but she can after she has changed this position (which she is free to do). The second premise (that if one is unaware of one’s sufficient reasons, one cannot act on them) is, in the relevant sense, false.

The situationist challenge has not ended, however. Consider the following case. A person walks by a dumpster. Unbeknownst to her, someone is inside the dumpster in desperate need of medical attention. The person, however, has no reason to suspect
this is so, and carries on walking. Assuming that there is a sufficient reason for her to help the person, is she free to do so? Well, she is free to open the dumpster and check for people in medical need, and, if she so acts, she is then free to help the person. However, she is free to help the person only if she checks the dumpster. She can’t help the person precisely because she is unaware that the person is in need. Thus ignorance of one’s sufficient reasons can indeed, in certain cases, rid one of the ability to act on one’s sufficient reasons. Perhaps the situational factors rid us of abilities to act on sufficient reasons in the same way the person’s being hidden in a dumpster does.

Why does the fact that the person in medical need is hidden in the dumpster prevent the agent from being able to act on her sufficient reasons? An initial answer is that this fact prevents the agent from knowing what her sufficient reasons are. But this is not the entire story. For imagine she knew that by searching the dumpster she would discover what her sufficient reasons are. In that case, she would be indeed be free to act on her sufficient reasons. A fuller explanation of why she cannot act on her sufficient reasons seems to be that she is unaware both of her sufficient reasons and of how to discover them. She is even unaware that there are hidden sufficient reasons around to be discovered.

We may then, present the following argument that some agents in situationist experiments are not free to act on their sufficient reasons:

1. There are agents in situationist experiments who (a) do not know their sufficient reasons, (b) do not know that there are such sufficient reasons of which they are unaware, and (c) do not know how to figure out their sufficient reasons (because of the specific circumstances they find themselves in).

2. If (a), (b) and (c) apply to an agent in such an experiment, this agent is not free to act on her sufficient reasons.

Some might object that, in this case, the agent does not have a sufficient reason, or any reason, to help the person in need. After all, she is entirely blameless for not helping, and she can justify her omission on the basis that she had no way of knowing about the person in the dumpster. However, there is surely some sense in which she has a reason to help. By performing certain actions (which are within her capabilities), she could help someone in desperate need. We might say that she objectively ought to help the person—doing so is what is objectively best. Another tempting way to put the point is that there is a sufficient reason for her to help, even if she does not have this reason.
3. Therefore, some agents in situationist experiments are not free to act on their sufficient reasons (because of their situations).

What should we make of this argument? It is better (though much more limited in scope) than the argument with which we started. The subjects who, due to situational factors, fail to realize that (for example) there is a person in need of medical attention nearby, or that a potentially dangerous event is unfolding, may well also be unaware of the need to look out for such facts, and, indeed, fail to know how to discover their sufficient reasons. Indeed, it is extremely plausible that this is so. Subjects, who, for instance, are in such a hurry that they do not realize that a person nearby them is in medical need, or who are caused by the presence of passive bystanders to interpret their situation as involving no emergency situation, are not simply unaware of their sufficient reasons, but also unaware of the fact that they are ignorant of their sufficient reasons. Indeed, the evidence that would provide them with knowledge of their sufficient reasons (were they to attend to it in the correct manner), is the same evidence that would provide them with the knowledge that they are ignorant of their sufficient reasons. Similarly, those subjects who interpret the smoke in a non-dangerous way are ignorant of the fact that there are sufficient reasons to act of which they are unaware.

Given this, such subjects also fail to know how to figure out their sufficient reasons. We must treat this claim carefully. Just as we may distinguish general ability from specific (freedom-level) ability, we may make the same distinction regarding knowledge-how. Someone may in general know how to figure out their sufficient reasons (attend to their evidence carefully, reflect on whether they might be missing something, etc.) without knowing on a particular occasion how to figure out their sufficient reasons. If a subject does not even suspect that they are ignorant of their sufficient reasons (and, indeed, may actively believe that they know them full well), then he is not in a position to know how to figure his reasons out. Though he may have the general cognitive capacities to work out his reasons, he does not know how to direct these capacities in his current situation.

It is very plausible, then, that there are certain subjects, influenced by certain situational factors, who are unaware of their sufficient reasons, or that there are such reasons to be discovered, and these subjects further do not know how to discover their sufficient reasons. Given premise 2, any such agents are not free to act on their sufficient reasons.

Premise 2 is hard to dispute. How might a person be perfectly free to act on her sufficient reasons without even knowing these reasons and without knowing...
how to find out what those reasons are? As far as we can tell, the only viable way of questioning premise 2 is the following—an agent may be free to act on her sufficient reasons even when she neither knows them, nor how to discover them, if she *does* know how to come to know how to discover them. Consider an analogy: an agent may be free to unlock a box even if she does not know how to open it, provided that she knows how to come to know how to unlock the box (perhaps she knows that, by reading the provided instructions, she will come to know how to unlock the box). We may even iterate this line of thinking: an agent may be free to A without knowing how to A, nor how to come to know how to A, as long as she knows how to come to know how to come to know how to A. We might thus question premise 2 as follows: even if subjects do not know their sufficient reasons, nor how to discover them, may they not still be free to act on their sufficient reasons if they *do* know how to *come to know* how to discover their sufficient reasons?

In answer to the question we may admit that if the relevant subjects know how to come to know how to discover their sufficient reasons, they are free to act on their sufficient reasons (though, it must be said, we do not feel compelled to admit this contentious claim). Premise 2 remains true, however, because exactly the same considerations that rule out the subject’s *knowing how to discover* their sufficient reasons also rule out the subject’s *knowing how to come to know how to discover* their sufficient reasons (etc.). The subjects fail to know how to discover their sufficient reasons in part because they are unaware that they are ignorant of their sufficient reasons. Accordingly, such subjects also fail to know how to come to know how to discover their sufficient reasons, as having such know how also involves (at the very least) knowing that there are sufficient reasons out there to discover. Thus premise 2 stands—if a subject does not know their sufficient reasons, nor how to discover them, she is not free to act on them. We may conclude that there are indeed agents whose (freedom-level) ability to act on their sufficient reasons is stripped by certain situational factors. Indeed, we may generalize our conclusion to agents outside of situationist experiments—*any* situational factor that prevents an agent’s knowing that she is ignorant of her sufficient reasons also prevents her from knowing how to discover such reasons (and from knowing how to come to know how to discover them, etc.), and thus removes her (freedom-level) ability to act on such reasons.

Proponents of this argument can concede that many situational factors do not strip us of our awareness of our sufficient reasons, or our abilities to act on them. They merely think that some agents (perhaps a small minority) are rendered unable to act on their sufficient reasons. This raises the question of how subjects who are not
stripped of their (freedom-level) ability are affected by their situations. In the following section we shall suggest that situational factors often make it *more difficult* for agents influenced by situational factors to act on their sufficient reasons than it is for agents not so influenced (all other things being equal).

5. Do Situational Factors Make Abilities Harder to Exercise?

Above we argued that those situational factors highlighted in situationist experiments sometimes eliminate agents’ freedom-level abilities. However, our argument applies only to a subset of those subjects adversely affected by situational factors. What of those agents who retain their freedom to act on their sufficient reasons, but who, nevertheless, do *not* act on their sufficient reasons? We propose that the data suggest that such agents find it *more difficult to exercise* their ability to act on various sufficient reasons. Take, as our test case, the bystander experiments. First, many more people in the “alone” condition help than those who are part of a group of observers. Also, given the subjects are assigned to their experimental condition on a random basis, it is likely that if we were to swap the experimental conditions for the subjects, we would still get somewhat similar results. This suggests that most of those people who fail to help the person in need of medical assistance or fail to report the smoke would intervene in the “alone” condition. Given that many of the subjects in each condition retain the relevant freedom-level abilities to act on their sufficient reasons, and assuming that all subjects share their sufficient reasons (which, by design, they *do* in the aforementioned experiments), it is notable that the vast majority of “alone” subjects act on their sufficient reasons and the vast majority of “accompanied” subjects do not. If the difference in situational cues does not alter the groups’ sufficient reasons, nor render most of one group unable to act on such reasons, nor, indeed, alter the longstanding moral values of either group, then we cannot account for the differences in behavior by pointing to any of these properties of the agents. Indeed, given that the majority of subjects in one group does help, while the majority in the other group doesn’t, and that we cannot point to a moral difference, or a difference regarding the abilities subjects possess between these groups (either or both of which could explain the difference in the majority behavior between the groups), it is very plausible that the relevant situational factor (presence of other people) makes it *more difficult* for people to exercise their ability to respond to their sufficient reasons (to help/react to an emergency). If the experimental subjects would, most likely, act on sufficient reasons in the absence of the pertinent situational cues (as indicated by
the data), and they have the ability (and are generally willing) to act on the relevant sufficient reasons, this suggests that something in their current situation obstructs this ability.

If a situational factor obstructs an ability without eliminating it, the factor makes the ability more difficult to exercise. But how should we understand the idea that an ability is harder to exercise in one situation than in another? The answer to which we are attracted appeals to the thought that, in one situation, more mental effort is required from the agent to exercise this ability than is required in the other situation. A situational factor makes it harder for us to exercise our abilities to act on sufficient reasons, then, if the situational factor influences us in such a way that we must expend greater mental effort in order to exercise these abilities.

What is mental effort? As we understand it, expending mental effort amounts to mobilizing energy for the purpose of meeting either cognitive goals (Gaillard (1993)) or executive goals. Cognitive goals relate to attaining knowledge, learning and comprehension (e.g., understanding a passage of text). Executive goals relate to formulating and carrying out action plans (e.g., decision-making, deliberation, resisting temptation, etc.). Mental effort includes both “task effort”, which needs to be invested in response to the computational demands of a given task, and “state effort”, which one needs to expend in order to shield one’s performance from potentially disrupting factors, such as fatigue (Mulder (1986); see also Fairclough & Houston (2004), Fairclough & Mulder (2012)). How much mental effort an agent needs to expend to meet any individual goal thus depends on the interplay of a variety of factors, including those internal to the agent (such as mood, desires, intentions, emotions, skills, psychological afflictions, etc.) and those external to the agent (for example, time pressure, task complexity, disrupting external factors such as noise and other distractions). When we speak of one’s ability as being harder to exercise, then, we can understand it in the following way. If agents A and B both have the same (relevant) abilities, it is harder for A to exercise her abilities than it is for B to exercise his, just in case it requires more mental effort from A to exercise her abilities than it does B. Let’s say Alice and David both possess the ability to suppress laughter in an inappropriate situation. However, let us also assume that is harder for Alice to employ or exercise this ability (and suppress her laughter) than it is for David (perhaps because Alice has problems with self-control generally). We suggest that Alice must expend more mental effort than David in order to exercise her ability—she must fend off greater distractions, attend to serious matters with sharper focus, close her mind to funnier thoughts, etc.
In other words, exercising this ability is more taxing on Alice—doing so drains more of Alice’s (mental) resources.

We may account for the situational data, then, by positing that certain situational factors give rise to various obstacles (e.g., perceived peer pressure, desire to conform or to obey the authority figure, etc.) to the agents’ acting on their sufficient reasons. While these obstacles are not (always) insuperable, it takes mental effort to overcome them—an amount of mental effort, indeed, that would not be required were those obstacles absent. There are, roughly-speaking, two relevant kinds of abilities to respond to sufficient reasons which might be adversely affected by situational factors—an agent’s epistemic abilities to recognize these reasons, and her actional abilities to translate such recognition into action.

Take the latter kind of abilities first. Consider a subject in the Milgram experiment who is aware of his reasons to refrain from increasing the level of shock the person in the other room (apparently) receives. While, we argue, he retains his freedom-level ability to thus refrain, the fact that an authority figure insists he continues with the experiment makes it very difficult to exercise this ability, and thus act on his sufficient reasons. He may, for example, feel considerable pressure to obey the experimenter’s commands; he may be somewhat overwhelmed by the novelty of his situation; he may feel less responsible for his actions, as someone else is taking charge. These facts are obstacles to him acting on his moral obligations even though he is aware of his moral obligations. This means that refraining from further shocking takes more mental effort—he must not only expend the effort needed to refrain from shocking someone period (i.e., the amount of effort he would need to expend were there no one ordering him), but he must also put effort into defying an authority figure, readjusting himself to the novelty of his situation, and taking responsibility for his actions. Exercising the relevant actional abilities thus may be more difficult when one is faced with pertinent situational factors.

The same applies to the former (epistemic) kinds of abilities which concern agents’ recognition of reasons. In the previous section we argued that if an agent is unaware both of her sufficient reasons and that there are sufficient reasons of which she is unaware, she does not know how to discover such reasons, and is thus not free to act on them. There may be cases, however, in which an agent is unaware of her sufficient reasons, but due to being presented with certain evidence, comes to believe that there

---

9. See Burger 2009 for explanations along these lines (and others) as to why people increase the level of shock so high in the obedience experiments.
are (or might be) sufficient reasons of which she is unaware. In such a case, the agent may know how to discover these reasons and thus be free to act on them. Even so, situational factors may make it more difficult for her to discover such reasons. Consider the seminarian who walks past the prone figure because he is in a hurry to deliver a talk. Suppose this particular seminarian is not aware that he should help the person in medical need, but is aware (at least in some dim way) that something is afoot that may warrant further attention. In such a case, the situational factor (in this case, being in a hurry) may still be an obstacle to the subject exercising his ability to recognize reasons. It is easy to see why this might be—the fact that he is in a hurry focusses his attention on other matters; his agreement to arrive on time may provide him with a sense of obligation to do so, which may in turn lead him to ignore other evidence concerning what he ought to do. Again, such obstacles to recognizing his sufficient reason to help are not insurmountable. Rather, it would take more mental effort than usual for the agent to recognize his sufficient reasons—he must not only interpret the evidence before him as providing a sufficient reason to act (as must anyone), he must divert his attention away from his current task, and reassess what he takes to be his pressing obligation at the time. All of this may take considerable mental effort.

6. Consequences for Moral Responsibility

How do the above considerations impact moral responsibility? A plausible line is this—subjects who are aware of their sufficient reasons for action, but do not act on them due to the influence of certain situational factors, are still morally responsible for their actions. After all, they know that they are not acting on their sufficient reasons, yet they are free to so act. That said, because acting on these sufficient reasons is harder, more effortful, because of the presence of situational factors, these subjects are thus less responsible (in particular, less blameworthy) than subjects who fail to act on their sufficient reasons without being influenced by adverse situational factors (other things being equal). Perhaps something similar is true of those subjects who are unaware of their sufficient reasons (because of certain situational factors), but are (perhaps dimly) aware that something is afoot. First, they are still morally responsible for their actions. After all, they fail to act on their sufficient reasons despite being able to seek out and discover such reasons (and then act on them). Second, they are less morally responsible for their actions than those not subject to situational factors. This is because it is harder, more effortful, for them to recognize their reasons than someone not subject to the relevant situational factors (again, keeping everything else equal).
Situational factors can thus be obstacles to both one’s acting on one’s (recognized) reasons and one’s recognizing reasons, and as such can reduce one’s responsibility in similar ways.\footnote{One’s responsibility when under the influence of situational factors is arguably even further reduced since one is not typically responsible for the circumstances (and thus situational factors) with which one is presented. Consider a case where an agent is free to choose whether she is in a situation in which it is relatively easy for her to do the right thing, or a situation where doing what one should is difficult. If she knowingly chooses the latter, and fails to do the right thing, then she is more responsible than someone who fails to do the right thing due to equally adverse situational factors which he did not choose.}

What of those agents who, due to situational factors, are unaware of their sufficient reasons \textit{and} of the fact they are out there to be discovered, and thus do not know how to discover or act on their sufficient reasons? We argued in 4.2.2 that such agents are not free to act on their sufficient reasons. Similarly, we think that such agents are not \textit{directly morally responsible} for failing to act on their sufficient reasons (by “directly morally responsible” we mean, roughly, responsible in such a way that this responsibility does not rest on the agent’s responsibility for earlier actions). One way to argue for this is by appealing to the following principle: an agent is (directly) responsible for failing to do something only if she could have done it (i.e., only if she was free to do it). Though this principle resembles the notorious and controversial Principle of Alternative Possibilities (an agent is morally responsible for doing something only if she could have done otherwise), the philosophical consensus is that (something like) the former principle is far more plausible than the latter (see van Inwagen 1983, Sartorio 2016). Still, we do not need to appeal to either of these principles. Indeed, it is already overwhelming plausible that agents who are not free to act on their sufficient reasons because they do not know them nor even how to discover them are not responsible for failing to act on them. Examples further confirm this. The agent who unknowingly passes the dumpster which contains a person in medical need is neither free to act on her sufficient reason to help the person nor directly responsible for failing to help. Similarly, a blind person who obliviously passes a prone figure without aiding the person is not free to help, nor directly responsible for failing to help. The lesson generalizes (more cautiously—we see no reason why it does not generalize): agents who are rendered unaware both of their sufficient reasons, and how to discover them, are not only unable to act on sufficient reasons, they are also not (directly) responsible for failing to act on their sufficient reasons.

There is a case to be made, however, that such agents may yet be \textit{indirectly} responsible for failing to act on their sufficient reasons. The seminarian in a hurry
who fails to process the situation as one in which someone needs help, or even one
in which further investigation is necessary might still be blameworthy for failing to
help because part of the explanation of why he fails in these ways is that he weighs
too heavily his reasons not to be late. Such weighing leads him to fail to be sensitive
enough to other reasons he might encounter which outweigh his reasons not to be
late. His failure to reflect on his reasons, or even see the need to, when faced with the
prone figure may thus stem from his earlier judgments and actions. These judgments
and actions involve a lack of attention to other normative reasons for action. In short,
then, though it is true that the seminarians do not recognize their sufficient reasons to
help, nor their reasons to reflect on their reasons, these facts do not necessarily render
them blameless for failing to help. This is because it is they who weight the importance
of not being late very highly, shift their attention from other moral considerations,
dismiss too quickly courses of action that diverge from their plan to arrive on time,
etc. Their epistemic failings, like their actional failings, are their own. It is still possible,
then, to hold the hurrying seminarians who fail to help indirectly responsible (indeed,
blameworthy) for failing to do so.\textsuperscript{11}

A brief summary is in order. We have argued that agents presented with adverse
situational factors, but who are nonetheless free to act on/discover their sufficient
reasons, are less responsible for failing to act on their sufficient reasons than they
otherwise would be, precisely because it is more difficult for them to act on such
reasons. We have also argued that those agents who are not free to act on their
sufficient reasons are not directly responsible for failing to do so, but may still be
indirectly responsible at least in part because they are responsible for the judgements
and actions that led to their being so unaware.

7. A Happy(ish) Ending

Above we discuss how different situational cues may hinder exercising our abilities
to act on sufficient reasons. Not all situational factors may impact such abilities
negatively though. Some studies show that situational factors may, in fact, make them
easier to exercise. Consider, for instance, a study by Bateson et al. (2006) in which the
experimenters tracked the amount of ‘honesty box’ contributions for refreshments in
relation to the type of picture presented on the instruction sheet placed above the

\textsuperscript{11}. This said, the way in which situational factors influence people seems relatively universal (i.e., they
influence the majority of us in similar ways). Perhaps this fact speaks in favor of further mitigating our
seminarians.
honesty box. People contributed to the honesty box, on average, 2.76 times more in those weeks when the information sheet had a picture of a pair of eyes rather than a picture of flowers. Given the results, it seems that being exposed to the images of eyes had significant influence on whether people acted on their sufficient reason (not to steal). Comparable results were found in a littering study by Ernest-Jones et al. (2011), where the odds of littering were halved in the presence of posters containing images of eyes, as opposed to posters containing images of flowers. Here too, arguably, situational factors (images of eyes) positively impacted people’s ability to act on their sufficient reasons (not to litter).

Such positive effects have been observed even in experiments which did not feature sufficient reasons. For example, Baron (1997) tested, on a sample of passersby in a shopping mall, how pleasant smells affect helping behavior. Subjects were asked to provide change for $1 which gave them a reason to help—although they did not have a sufficient reason to help (providing change would be best described, in normal circumstances, as supererogatory).\(^\text{12}\) It was found that helping behavior was significantly greater in the presence of pleasant fragrances than in their absence. For instance, in one of the conditions, it was found that 60% of the women exposed to pleasant smells helped, as compared with only 16% of the women not subjected to such smells. This indicates that normatively irrelevant situational factors may sometimes be beneficial with regards to our ability to act on supererogatory reasons.

While, in some ways, it is a desirable result that situational factors may have a positive impact on our abilities to act on reasons (sufficient or supererogatory), the above data further illustrate just how sensitive such abilities may be to different circumstances (often without us being aware of them being so sensitive). This raises a general worry about agents consistently and reliably exercising these abilities, across different situations. From the point of view of moral responsibility, it is a cause for concern that the exercise of our abilities is (too) dependent on luck—in other words, it is undesirable that one’s abilities being easier or harder to exercise (significantly) depends on whether one (luckily) faces favorable situational factors.\(^\text{13}\)

To sum up, then: situational factors sometimes strip agents of their abilities to act on sufficient reasons, while situational factors often make it harder to exercise such abilities. This in turn suggests that agents thus influenced by situational factors are

\(^{12}\) Supererogatory actions are those above and beyond the call of duty. If you have a supererogatory reason to A, your Aing would be admirable, but not obligatory.

\(^{13}\) We explore the connection between situationism and moral luck in Herdova & Kearns 2015.
less morally responsible for their actions than those agents who are not so influenced. Lastly, however, it seems there are also situational factors which enhance the agents’ abilities to act on their sufficient reasons (and even on their supererogatory reasons)—such situational factors make it easier for agents to exercise their abilities to act on their reasons (sufficient or otherwise). There is some reason to worry, then, that we are unable to do the right thing given the subtle influence of our situations. We should be even more wary, however, that doing the right thing, and consistently so, might be more challenging than previously thought.

References


Nationalism & Social Stability

Stephen McAndrew
University at Buffalo

Biography
Stephen McAndrew is a graduate student in the philosophy PhD program at the University at Buffalo (SUNY) and also holds a JD. His research is concentrated on examining how economic transactions can reduce prejudice between diverse groups like immigrants and nonimmigrants.

Publication Details

Citation
Nationalism & Social Stability

Stephen McAndrew

Abstract
There is an increasing turn to nationalism around the world. The advocacy of “America First” policies, the Brexit leave campaign in Britain, and recent elections in Poland and Hungary show evidence of a rise in nationalistic sentiments. One reason given to explain this rise in nationalism is that in an increasingly diverse world stability is not possible without close cultural links between members of society, and that a shared national culture can provide those links. Nationalists argue that a shared national culture is a necessary condition of creating social solidarity that creates social stability. However, the nationalist solution to creating social solidarity can be questioned on a number of counts. First, it relies on a conceptually problematic account of national identity that holds that national identity includes elements like a shared culture, a shared history, and a shared connection to a particular geographic territory. There are reasons to think that this type of account of national identity is closer to an account of ethnicity than an account of national identity. An ethnic nationalism is morally problematic as it contends that one can only have solidarity with those who share one’s ethnicity, and could be used to justify discrimination against ethnic minorities. Second, even if this is the correct account of national identity, it is not the case that a shared culture, a shared history, and a shared connection to a particular geographic territory are necessary or sufficient conditions for social solidarity. Finally, nationalist attempts to protect the national identity of a liberal democracy by restricting all immigration, may actually destroy the values, such as individual rights and limited government of that liberal democracy.

Keywords
Nationalism, Social Solidarity, National Culture, Liberal Democracy, Immigration

Introduction
There is an increasing turn to nationalism around the world. The election of Donald Trump and his advocacy of “America First” policies, the Brexit leave campaign in Britain, and recent elections in Poland and Hungary show evidence of a rise in nationalistic sentiments. Regimes in Russia, Turkey, China, and India also show nationalistic tendencies (Fukuyama 2016; Economist 2016). One reason given to explain this rise in nationalism is that in an increasingly diverse world stability is not possible without close cultural links between members of society, and that a shared national culture can provide those links. Nationalists argue that a shared national culture is a necessary condition of creating social solidarity that creates social stability. For example, David Miller argues that “[p]eople feel emotionally attached to one another because they share this identity. They feel that they belong together and have responsibilities to
each other that are not simply the result of existing institutions and practices” (Miller 2016, 27). There is empirical evidence of a lack of solidarity between immigrants and nonimmigrants where there are cultural and religious differences. Philipe van Parijs points to the lack of integration of culturally diverse immigrants in Brussels (Van Parijs 2004) and Sweden is having a difficult time integrating immigrants (Traub 2016). So, nationalist concerns cannot be completely dismissed as without any empirical basis. It is also the case that it is not wrong to want social solidarity and hence social stability - no one wants to live in a society of constant conflict and tension. So, instead of dismissing nationalist worries about cultural diversity and social solidarity as motivated by misinformed or even malicious attitudes towards immigrants, I will take the nationalist concern in good faith and examine whether or not a robust national culture or identity is necessary for social stability.

To discuss nationality and how it is related to social solidarity I will examine an account that argues that a shared national identity is necessary for social stability. David Miller is a prominent proponent of this type of view so this paper will analyze his position. Miller argues that a national identity or nationality aids social solidarity, which (1) makes social policies like redistribution and environmental protection easier to implement, and (2) makes deliberative democracy easier because society will not collapse into sectional interests fighting each other (Miller 1995, 90–99). Miller defines national identity as communities: (a) constituted by belief by the members of the community that they form a nation, (b) that have historical continuity, (c) that are active in nature meaning that they do things like make decisions together, (d) that are connected to a particular geographical territory, and (e) that share a common public culture (Miller 1995, 27). Miller argues national identity is not directly linked to ethnicity because this will lead to racism. Miller states that nations are not solely ethnic and that a public culture is not necessarily based on ethnicity but may, for example, be based on the mixing together of many different ethnic groups. Miller contends that national identity helps to make people feel like they belong in a particular physical place. Because nations have historical continuity, Miller contends that members of a nation have an obligation to their forebears and those who come after them to continue in the same historical traditions and to maintain the nation that has been fought for. Miller acknowledges that historical continuity that forms part of national identity can have large mythical elements. It may be that the origins of the nation are lost in history and that stories that have contemporary relevance or utility are

---

1. Miller uses the terms national identity and nationality synonymously and I will do the same here.
projected back on to the past. It can also be the case that some historical events that are eulogized as evidencing the national culture may have many mythical elements. Miller holds that even though these myths of national foundation won’t stand up to critical examination, they are valuable for two reasons: (1) they provide reassurance that the national community we are part of has historical roots; and (2) they encourage us to live up to the moral example of our founders. For example, ‘the Dunkirk spirit’ encourages those with British national identity to persevere and work together to accomplish a great task (Miller 1995, 35–38). The public culture for Miller does the work of bringing people together by providing a sense that they belong together and have not been merely thrown together. Though a national culture is not monolithic or all-embracing it includes social norms like queuing. It is important to note that Miller is not against immigrants so long as they can be assimilated to the public culture, but he would seem to be committed to seeing it as problematic if immigrants come with a very different culture.

In Section A, I will look at whether or not Miller’s definition of nationalism is correct or if his definition of nationality is closer to that of ethnicity. If his thesis is that people need to share ethnicity to feel solidarity this will not work in a multiethnic context. So, it will be problematic for his account of national identity if it is more like an account of ethnicity. In Section B, I argue that even if Miller’s definition of national identity is correct, it is still the case that none of Miller’s elements of nationality unproblematically generate solidarity. In particular, I will examine whether historical continuity, having a connection to a geographic location, or having a shared culture are necessary or sufficient conditions for solidarity. Finally, in Section C, I argue that attempts to completely protect national identity, as defined by Miller, by restricting all immigration are likely to backfire and instead undermine national identity, particularly if a nation views itself as a liberal democracy.

A. Does Miller have the correct definition of Nationality?

Some theorists have questioned whether Miller’s definition of nationality is correct. Jorge Gracia contends that national identities are based in laws. Gracia argues that nations are political entities, governed by laws and tied to a territory and that membership of the nation is set by the laws established by the government of that nation (Gracia 2005, 109–141). Gracia rejects Miller’s definition of nationality for the following reasons. Gracia thinks that Miller’s requirements that nations are active in character and are constituted by the beliefs of the members of the nation are too
vague to serve as part of a definition of a nation (Gracia 2005, 110–111). Gracia argues that a shared public culture is not sufficient to create a nation because two separate nations may share the same public culture. He points to the example of a shared culture between parts of Argentina and Uruguay (Gracia 2005, 110–111). He argues that historical continuity is conceptually problematic as it leads to questions as to how long a history a nation must be able to attach to itself (Gracia 2005, 111). Gracia argues that “the unity of a nation does not involve features common to its members. Its unity should not be understood in terms of linguistic, cultural, racial, genetic, experiential, or class boundaries, even if in context these further unite the members of a nation” (Gracia 2005, 128). Gracia defines a nation as follows:

A nation is a subgroup of individual humans who satisfy the following conditions: they (1) reside in a territory; (2) are free and informed, and (3) have the common political will to live under a system of laws that (i) aims to ensure justice and the common good, regulating the organization, interrelations, and governance, and (ii) is not subordinated to any other system of laws within the territory in question. (Gracia 2005, 130)²

In fact, Miller’s definition of nationality is closer to what Gracia would define as an ethnic group. Gracia’s definition of an ethnic group is not based on descent but on kinship. That is to say that just as someone can marry into and so join a family, one can marry into or join an ethnic group. Gracia contends that historical relations between members of an ethnic group connect them to each other (Gracia 2005, 45–55). This is similar to Miller’s idea that historical relations between members of a nation bind them together. Gracia also points out that the origin of many ethnic groups may contain some element of myth. For example, he discusses the view of some ethnic Cubans of pre-Castro Cuba that contains mythical elements (Gracia 2005, 55–57). Again, this is similar to Miller’s conception of nationality, which Miller involves historical continuity and foundation myths. Gracia defines an ethnic group or ethnos as follows:

An ethnos is a subgroup of individual humans who satisfy the following conditions: (1) they belong to many generations; (2) they are organized as a family and break down into extended families; and (3) they are united through historical relations that produce features,

². Gracia holds that nationality is a relational property of belonging to a nation that is shared by members of a nation.
that, in context, serve (i) to identify the members of the group and (ii) to distinguish them from members of other groups. (Gracia 2005, 54)

However, historical continuity is not the only point of similarity between Miller’s definition of national identity and definitions of an ethnic group. Other theorists have argued that ethnicity depends on members of the ethnic group sharing a common culture or a common language. For example, Angelo Corlett argues that a number of factors go towards establishing an ethnic group including a common culture, and a common language (Corlett 2003, 11–13). Anthony Appiah (Appiah 1990, 498) and Michael Brown (Brown 1997, 81–82) also think that an ethnic group has a common culture. In fact, Brown’s definition of an ethnic group is very similar to Miller’s definition of national identity. Brown holds that an ethnic group consists of the following elements: (i) a group name, (ii) belief in a common ancestry, (iii) shared historical memories, (iv) a shared culture, (v) connection to a particular geographic territory, and (vi) the group must believe they constitute an ethnic group (Brown 1997, 81-81). Once again, we must consider whether it makes more sense that sharing a common culture is a feature of an ethnic group or a feature of a national identity.

So many of the elements that Miller thinks make up national identity are thought to be part of what makes an ethnic group by other theorists; in particular the requirements that a national identity have historical continuity and a common culture. Indeed some commentators have labeled the rise of populist nationalist movements in the past few years as “ethnic nationalism” (Economist 2016). So, there is a case that Miller is in fact advocating not a shared national identity is necessary for social solidarity but instead that shared ethnicity is necessary for social stability. This, as Miller recognizes, is problematic for his position.

Miller does not want to say that nationality boils down to ethnicity, and he explicitly denies that it does (Miller 1995, 19-21). Miller’s reason for denying that his definition of national identity is a definition of ethnic group is that Miller holds that his definition of national identity does not have a descent as a necessary condition (Miller 1995, 19-21). However, theorists like Gracia argue that this means that Miller’s definition is not based on race, but may still be based on ethnicity as Gracia argues that ethnicity is based on kinship, while race is based on descent (Gracia 2005, 85). So,

3. Gracia defines ethnicity as “the relational property of belonging that characterizes the member of an ethnos.” (Gracia 2005, 54)
stating that his account is not based on descent may not be enough to defend Miller’s account of national identity from charges that it is race-based ethnicity but not from charges that it is an ethnic based nationality.

An ethnic nationalism is problematic for two reasons. First, if social solidarity depends on everyone sharing the same ethnicity, solidarity will be difficult to achieve in the current world where most nations are composed of a multiplicity of different ethnic groups. This is a denial of addressing the problem of creating solidarity between diverse groups. Second, it is morally problematic to hold that one can only have solidarity with those who share one’s ethnicity. Ethnicity like race is not a morally relevant characteristic that can be permissibly used to treat people differently. If one argues that there can be no social solidarity in a multiracial or multiethnic society, this is not going to be a very convincing argument because this is an argument that can be used to support ethnic and racial discrimination.

However, given the controversy over the definition of national identity, it may be better to place less weight on the definition of national identity, and look to the factors that are thought to connect and build solidarity within groups. To be charitable, it should also be noted that even if Miller is mistaken in his analysis of national identity, he is still claiming that certain features that he includes as part of national identity are important to creating social solidarity and maintaining an open democracy. Moreover, it may also be the case that the factors that Miller thinks are elements of national identity are connected with national identity in the popular imagination of those who are seeking to use nationalism as a reason to restrict immigration. For these two reasons it is still a valuable exercise to examine the elements of Miller’s definition of national identity to see if they actually can create social solidarity necessary for political stability. If we take Miller’s list of features of nationality ((a) through (e)) above, the three that appear to do the most work are historical continuity, connection to a particular territory, and a public culture. If it can be shown that these features are neither necessary nor sufficient to create solidarity then Miller’s account, whether it is of nationality or ethnicity, will not have shown that these features are necessary to maintain social solidarity. This is the task I will turn to in the next section.

**B1. Having historical continuity does not necessarily generate solidarity**

Let us first examine whether or not historical continuity is necessary or sufficient to create social solidarity. Being tied to the past is not sufficient to create to social stability. Even when people are looking to the same historical events and founders, they can
come to hold very different ideas, even opposing ideas, and these opposing ideas can cause instability. If the same historical events can give rise to many different national identities, including some that conflict with each other, then historically grounded national identities do not seem to guarantee social stability or deliberative democracy.

Take for example, Patriotica, a medium-sized possible nation. In Patriotica, there is a story that its earliest inhabitants, who are historically linked with its present inhabitants, fought for freedom from the nearby and much larger state of Imperica. The story is very ancient and many of the details are vague and some of the specific stories contain a large element of myth. Everyone agrees on the fact that gaining freedom from Imperica involved a lot of violence and death as there is evidence of preserved remains of mass graves of mangled skeletons, whom genetic tests have shown to be related to current inhabitants. The myth of sacrifice for Patriotica is taken to have different meanings by two different groups. One group (the “Warriors”) holds that these stories of sacrifice mean that the national identity of Patriotica is militaristic and that physical strength and willingness to use violence to protect oneself or Patriotica are national traits of Patrioticans. The Warriors further hold that current Patrioticans owe their allegiance to past and future generations to maintain these militaristic traits. They owe allegiance to ancient Patrioticans who died to birth Patriotica, and to future generations to preserve Patriotica in its current form (or close to its current form).

However, another group (the “Peacemakers”) hold that these stories of sacrifice point to the need to devote one’s life to preservation of peace of Patriotica that was bought at such a great price of loss of human life. The Peacemakers also hold that they owe it to past and future Patrioticans to maintain these pacifist national traits. Both of these groups look to the same origin myths but come to very different conclusions. In fact, the different views of the Warriors and the Peacemakers have caused conflicts on many issues. The Warriors want to spend huge resources on military spending and to tailor school curriculum to train children to be tough and willing to die for Patriotica. The Peacemakers believe that resources should be devoted to healthcare and the provision of high-quality social services for all. They want military spending greatly reduced. They believe that the school curriculum should make children aware of the perils and costs of war and should focus on making children empathetic to all members of society. Because both groups are roughly equal in number they have relatively equal political power. However, each group believes that the other group is endangering not just the national identity of Patriotica but its very survival. The Warriors think that the Peacemakers will neglect the defense and security of Patriotica by not focusing on preparing the nation and the citizenry for war. The Peacemakers think that the
Warriors will lead Patriotica into unnecessary and costly conflicts that threaten the survival of the nation. Because of the relative equality of political power of both groups, neither can politically impose its view of Patriotica on the other group. This has led to frustration on both sides and to public protests and counter protests that increasingly result in violence and even loss of life. So, the two different worldviews are leading to instability and a lack of solidarity. Both worldviews are grounded on the historical origins of Patriotica but instead of bringing the society together, it tears it apart. It is plausible, as this thought experiment shows, that shared stories of national origin are not sufficient to generate social solidarity and prevent society from collapsing into conflicting sectional interests.

So, being tied to the past does not guarantee social stability because even when people are looking to the same historical events and founders, they can come to hold very different ideas, even opposing ideas and these ideas that can cause instability. It is plausible that the Patriotica thought experiment has parallels in the actual world. In the U.S., liberals and conservatives both appeal to founding fathers and founding documents and come to very different positions. Take for example debates over how to understand the U.S. constitution, which is a historical document related to the founding of the nation of the U.S. Liberals hold that the constitution is a living document that should evolve to protect rights not contemplated by its drafters, including same sex marriage and abortion. Whereas conservatives argue that the constitution should be narrowly interpreted in accordance with the actual text and not in line with changing social conditions. So, liberals and conservatives both look to the U.S. constitution as a historical founding document of their nation, but interpret it differently leading to a lack of social solidarity over issues like same-sex marriage and abortion. It may be the case that the same history can ground so many different national identities that the importance of being attached to the past becomes difficult to ascertain. If the same historical events can give rise to many different national identities, including some that conflict with each other, then historically grounded national identities do not seem to guarantee social stability or deliberative democracy.

It also seems that historical continuity is not necessary for social stability. Let us return to the origins of Patriotica. Imagine that it is a large island that is located a long way from any other significant land mass. Further, imagine a great storm drives sailors who originate from lands to the east, west, north, and south of Patriotica to run aground in the waters around the previously uninhabited Patriotica with no immediate hope of returning to where they came from. We can imagine these sailors, who have no previous historical relations, forming a stable society on Patriotica. That is not to
claim that they will inevitably form a stable society, but just the modest claim that a lack of historical relations will not stop them from doing so. Some features that create solidarity between the sailors from different parts will have to be operating but none of those features have to be historical continuity. One possible feature of creating such solidarity will be the need to cooperate with each other to survive. They may also need to pool and trade expertise and knowledge in order to survive. One might object that this will not create massive amounts of social solidarity. However, an account of social solidarity that creates stability only needs to create enough solidarity to create stability. It does not have to create blood brothers just partners. An empirical example of social solidarity coming about without the necessity of historical continuity is again the U.S. where the absence of long historical continuity did not prevent social solidarity and a stable society becoming established. So, historical continuity is neither sufficient nor necessary to create social solidarity and stability. Perhaps connection with a particular geographic territory is better for creating social solidarity.

**B2. Connection to a Particular Geographic Territory**

Is connection to a particular geographic territory necessary or sufficient to creating social solidarity? It seems that feeling that one belongs in a particular physical place is not sufficient to generate solidarity between others who also feel that they belong to the same geographical location. The sense that one belongs in a particular geographical place has been the source of much conflict. For example, Israelis and Palestinians feel connected to the same geographic area, but this shared connection is a source of conflict. The same can be said of Catholics and Protestants in Northern Ireland and in many other places. Therefore, there must be something more than feeling like one belongs in a particular geographic region that generates solidarity.

However, is connection with a particular territory necessary to create social stability? To return to the example of the sailors who find themselves shipwrecked on the island of Patriotica, it seems perfectly plausible that they could create solidarity between each other even though none of them have a particular connection to the territory of Patriotica. They may even feel connected to the territories of the countries from which they originated. Again, if they create solidarity between themselves it will be due to factors other than feeling a connection to Patriotica. Immigrants came from many different places over the world to the U.S., for example, from Germany, Italy, Ireland, England and had no connection to the territory of the U.S., but were able to
generate social solidarity. These immigrants may have felt deep connections to their ‘homelands’ but this was not a barrier to the generation of social solidarity in the U.S.

So, it seems that connection to a particular geographic territory is neither necessary nor sufficient to creating solidarity. However, perhaps sharing a public culture is something that can create social stability.

**B3. Public Culture**

Miller argues that citizens have a right to their national identity that entitles them to control how their national culture is changed. Because the entry of people with diverse social and religious identities might change that culture, Miller holds that citizens have a right to control immigration in order to protect their national culture. Miller will allow immigrants to enter so long as they assimilate to the national public culture, but then the worry is that some groups will be seen as having religious or cultural commitments that prevent them from assimilating. Do people have a right to protect their national culture? It seems that Miller can only claim this if the public culture creates solidarity that would be lacking with it, so that depriving a group of their public culture would deprive them of a stable society. However, if public culture is neither necessary nor sufficient to creating stability, then the stakes are not so high and this argument has less force. So, in this section I will start by examining whether or not a shared public culture is necessary or sufficient for social solidarity.

Is a shared national public culture necessary or sufficient to create social stability? It seems again that the shipwrecked sailors on Patriotica can be used to show that sharing a public culture is not necessary to social stability. They arrive there from different places and it seems plausible that they can create stability without a shared culture. They will need rules to govern their interactions, and if that is what is meant by culture, there will be some form of culture. However, will they need common forms of music, art, manners, etc. to live in a stable society? It seems plausible that they would not. If the sailors remain on Patriotica and settle there and have descendents then a common culture will likely develop, but in that scenario the common culture comes out of a framework that facilitates social stability and is not a condition of such a framework. So, conceptually it is not a necessary condition to social stability to share a common culture. Historically, countries like the U.S. and Canada have been able to maintain social solidarity in the face of immigrants arriving from many different cultures.
But is a national public culture sufficient to create social solidarity? Well, first of all we would have to figure out what a public culture is. Miller is not clear, and this may be because it is very hard to get to specifics regarding what a public culture is. He writes that a public culture is not “monolithic and all-embracing” (Miller 1995, 26). He mentions that it might include queuing, filling in forms honestly, and a commitment to democracy or the rule of law and might extend to religious beliefs or a commitment to preserve the purity of the national language (Miller 1995, 26). Given this characterization of public culture it is conceivable that a shared national public culture may not in every case maintain social solidarity. We could imagine that Patriotica has a strong public culture of shared manners, language, food and a strong public culture when it comes to what constitutes great art but still has deep political differences between the Warriors and the Peacemakers as described in Section B1 above. In the U.S., even though liberals and conservatives share the same language and enjoy the same food there are increasingly deeper political differences between the two groups.

However, as I mentioned in the introduction, there are empirical examples of there being cultural differences between immigrants and nonimmigrants and a lack of integration of those immigrants in places like Belgium and Sweden. So, while sharing a public culture may not be necessary or sufficient to create social solidarity, the lack of a shared public culture seems to be related to a lack of social solidarity in these cases. So, what can be said about such empirical examples? However, to the extent that those who are worried about a lack of integration between immigrants and existing citizens where there are big cultural and religious differences between immigrants and nonimmigrants, we should keep in mind the differences between the following propositions:

1. Cultural differences can cause a lack of social solidarity.
2. Sharing the same culture is necessary for social solidarity.

The argument that David Miller wants to make based on the scenario we observe cultural differences and a lack of social solidarity is the following:

1. Cultural differences can cause a lack of social solidarity.
2. Immigrants come to their new country with cultural differences from the existing population.
McAndrew

3. In order to create social solidarity between immigrants and existing citizens, the immigrants must assume the culture of the existing population. (From 1 & 2).

4. Sharing the same culture is necessary for social solidarity. (From 3).

However, the conclusion (3) does not necessarily follow from premises (1) and (2). What follows from premises (1) and (2) is instead:

3.* Steps should be taken to create solidarity between immigrants and existing citizens.

What steps consistent with (3*) can be taken to create solidarity between immigrants and existing citizens? David Miller wonders if there will be divergent conceptions of social justice and a lack of trust in a multicultural society. He also worries will those in a multicultural society be willing to extend principles of social justice to those outside of their group (Miller 2004). However, this does not show that immigrants have to take on the culture of their new country in order to create social solidarity necessary for political stability. In fact, Miller recognizes that the problem is not so much due to cultural differences but a lack of contact between cultures. He describes the amount of trust and the level of shared concept of social justice in a diverse society in three situations. The first in which the groups are alienated from each other with little contact even though they live alongside one another. The second in which the groups are segregated, and the third in which groups are integrated “and interact in cross-cutting associations of various sorts” (Miller 2004). Miller argues that this third scenario generates generalized trust and a shared sense of social justice. Similarly, Van Parijs notes that increased contact between groups tends to create greater economic solidarity (Van Parijs 2004). In fact he suggests that bringing immigrants into contact with existing citizens will bring groups together. This suggests that some mechanisms should be in place to encourage contact between immigrants and the existing population because that will help create solidarity. Interestingly, Van Parijs speculates that increased contact between immigrants and existing population will not adversely affect the culture of the existing population as much as it will adversely affect, and substantially change the culture of the immigrant population. So, increased contact between immigrants and existing citizens would seem to be a step taken to increase solidarity between immigrants and existing citizens.
The next issue is how this increased contact can be brought about. Miller argues that in order to be fully assimilated into their new location, immigrants must be culturally integrated and not merely politically and economically integrated. Miller argues that sharing cultural values is necessary to create trust, which trust is necessary for political and economic integration (Miller 2004, 29). But a case can be made that economic integration can build trust and lead to social integration. If we are permitted to return to the Patriotica thought experiment with the shipwrecked sailors from different origins we can show how trust can be gained via economic transactions. They may have different skills and need to trade with each other in order to survive. If they are to trade successfully they will need to trust each other. To create economic contact between immigrants and nonimmigrants policies might be put in place to facilitate immigrant business start-ups in a facility like a bazaar where immigrants can sell goods or offer services that will bring immigrants into contact with existing citizens.

So, while there are empirical cases where there are cultural differences between immigrants and nonimmigrants and a lack of immigrant integration, such cases do not necessarily support the contention that immigrants must take on the culture of their new country in order for there to be solidarity. This can be seen from the fact that (4) does not follow from (3*). But, these empirical cases demand an explanation. Why is it the case that cultural diversity between immigrants and nonimmigrants can be observed along with a lack of immigrant integration in certain cases but not others? We should be careful to make sure that the problem of lack of immigrant integration is not due instead to multicultural policies that can prevent immigrants from assimilating with the existing population. However, the kind of multicultural institutions that build solidarity between diverse groups is not a subject that I have space to address here. I briefly suggested policies that facilitate immigrant business ownership that brings immigrants and nonimmigrants into greater contact, but again I do not have the space to explore this idea fully here. All I have sought to show is that just because cultural differences can lead to a lack of social solidarity it does not follow that sharing the same culture is necessary for social solidarity. However, in Section C I want to examine another problem with national identity as a protection for social and political stability, that it may achieve stability but at a high cost for liberal democracies.
C. Harming National Identity By Trying to Protect it

Miller argues that shared national identity is necessary to generate the solidarity needed for what he calls an open, deliberative democracy to function.\(^4\) In Section B, I contended that national identity, at least of the kind defended by Miller, is neither necessary nor sufficient for social solidarity that creates political stability. In addition nationalism that seeks to protect a liberal democracy by preventing those with different cultural or religious values from entering, may actually be extremely harmful to that liberal democracy. This is because the policy measures necessary to retreat to a robust national identity that excludes those who are a cultural threat to that national identity are themselves a threat to a national identity built on liberal democracy.

Let us return again to Patriotica. Patriotica is faced with numbers of culturally diverse people who wish to live in Patriotica. Nationalists in Patriotica argue that allowing such people to live in the country poses a threat to the country’s national identity and therefore a threat to its stability. Patriotica is also a liberal democracy and this is part of its national identity. By liberal democracy I mean that Patriotica is a society in which: (i) the citizens enjoy rights like freedom of association, freedom of expression, freedom of movement, and due process, and (ii) the government is limited by the rights of citizens. My argument is that if a retreat to national identity means that Patriotica must close its borders to all culturally diverse people, then it may not be possible for Patriotica to remain as a liberal democracy.

Enforcing a complete ban on immigration will require force to prevent immigrants from entering Patriotica. Government agents will have to use force and the threat of force to keep immigrants from entering the territory of Patriotica. If immigrants try to evade the authorities of Patriotica, then force will have to be used. This may mean use of force not just at official border crossings but at other potential points of entry in the country. Controls on entry will require citizens of Patriotica to be checked as they enter. These measures may not unduly worry citizens of Patriotica or affect the status of Patriotica as a liberal democracy. However, it is likely that if the goal is the complete elimination of immigration that more will be needed to be done as some immigrants will arrive into Patriotica.

In order to locate and remove all immigrants that enter through unofficial channels, government agents will have to investigate businesses and residences to check the

\(^4\) Miller defines a deliberative democracy as “the ideal of a political community in which decisions are reached through an open and uncoerced discussion of the issue at stake where the aim of all participants is to arrive at an agreed judgment” (Miller 1995, 96).
credentials of those living and working there. Inevitably, some citizens of Patriotica will have their credentials checked in their places of work and business. Having to prove your right as a citizen of Patriotica to reside there is not something that one envisions as part of a citizen of a liberal democracy. It is further not part of a liberal democracy to just target certain people due to their names or physical appearance as likely immigrants. A liberal democracy treats all of its citizens equally and cannot differentiate along racial or ethnic lines. Keep in mind that national identity can include those of different ethnicities. So, all citizens of Patriotica will experience citizenship checks and have their privacy from government intrusion curtailed. Rules punishing landlords from renting to immigrants or punishing employers from hiring immigrants are not likely to be enough to stop the citizenship checks. First, the government will need to conduct checks to make sure that landlords and employers are complying with the rules. Second, in a non-ideal world some landlords and employers will attempt to evade the rules.

To completely eradicate all immigration to protect national identity would further require random public stops of pedestrians and motor vehicles. Stops on public transport could also be used to locate immigrants. Stops on public and private transportation would be a restriction on one’s freedom on movement that does not seem to belong in a liberal democracy. Again, it would not be compatible with a liberal democracy to restrict those stops only to those suspected of not being citizens of Patriotica due to their name or physical appearances.

In a liberal democracy many people may not only oppose these government intrusions in their lives, but oppose the policy of restricting immigration. In Patriotica the Peacemakers may embrace newcomers seeing it as an opportunity to broaden their society and practice empathy with those different from themselves. Any views opposing the restrictive immigration regime will have to be kept from being advocated. This will require restrictions on freedom of expression and freedom of association of the Peacemakers. Again, this will be problematic for a liberal democracy in which issues like immigration are openly and freely considered.

The changes necessary to insulate a society like Patriotica from all unofficial immigration would significantly affect Patriotica as a liberal democracy. And, therefore, adopting these changes would change what it means to be a citizen of Patriotica, if part of Patriotican national identity is being part of a liberal democracy. The overall point is that in order to stop all immigration, Patriotica will have to turn into a police state restricting rights of Patriotican citizens. This will affect the status of Patriotica as a liberal democracy, and change what it means to be Patriotican. It is likely that such
changes will be seen as a greater threat to what it means to be a Patriotican than the fact that some unofficial immigration will occur. Just like a soccer player who in seeking to defend his team’s goal mistimes his kick and scores in his own team’s goal, the steps to protect a liberal democracy from unofficial immigration would result in an ‘own goal.’ It is therefore my thesis that completely insulating a country like Patriotica from all who want to enter is not feasible without severe damage to Patriotican national identity as a liberal democracy. I think that any liberal democracy faces this same issue - that by insulating itself from all migration that it creates a greater threat to its cultural identity than migration poses.

Someone might object that liberal democracies like the U.S. or the U.K. do not actually protect individual rights consistently, so being a liberal democracy cannot be part of their national identity. However, national identity does not have to be based on a completely accurate picture of a nation. Instead, it is based on the way the citizens of that nation view it. So, it can be the case that being a liberal democracy is part of the national identity of a nation even if individual rights are not protected in all cases so long as the citizens view being a liberal democracy as part of the identity of their nation.

**Conclusion**

The nationalist solution to creating social solidarity can be questioned on a number of counts. First, it relies on a conceptually problematic account of nationalism that holds that national identity includes elements like a shared culture, a shared history, and a shared connection to a particular geographic territory. There are reasons to think that this type of account of national identity is closer to an account of ethnicity than an account of national identity. Second, even if this is the correct account of national identity, it does not seem to be the case that a shared culture, a shared history, and a shared connection to a particular geographic territory are necessary or sufficient conditions for social solidarity. It may be the case that immigrants have different cultural values and that some mechanisms may have to be put in place to integrate, in particular by bringing immigrants into contact with the existing population, but this does not show that a shared public culture is necessary for social solidarity. Finally, nationalist attempts to protect national identity of a liberal democracy by restricting all immigration, particularly in a liberal democracy, may destroy the values, such as individual rights and a limited government of that liberal democracy.
References


Are There Benefits to Benefits?

Holly Stevenson
The University of Iowa

Biography
Holly is a third year PhD student in Philosophy at The University of Iowa and has received her master’s degrees in Philosophy and Gender, Women and Sexuality Studies from the same institution. Holly received a 1st Class Honors Degree (MA) in Philosophy and the McClaggan Prize in Moral Philosophy at Senior Honors Level from The University of Glasgow, Scotland.

Publication Details

Citation
Are There Benefits to Benefits?

Holly Stevenson

Abstract
The concept of benefits has long been at the heart of discussion in political philosophy. Many political philosophers, both contemporarily and historically, have used benefits to ground political obligation and subsequent theories which stem from political obligation. Many philosophers claim that the government having provided us with “benefits” grounds our political obligation to obey the laws of the state. Benefits are often used as grounding for political obligation which then stems into fair-play theories, consent theories, gratitude theories, associative theories etc. Thus, it is clear that benefits have a key role in the fundamental justification for political obligation and the subsequent theories which arise from it. It is important at this stage to outline what the aims of my paper are and what they are not. Firstly, my paper is not arguing for any one viewpoint. I have no commitments to whether we should incorporate benefits into our political theorizing. What is clear to me is that benefits are very much a part of political philosophy and that is why they are the subject of my paper. Similarly, I am not arguing for or against political obligation or any theories which incorporate it. However, the link between benefits and political obligation justifications is well-documented in the literature, and therefore requires philosophical investigation independently of whether discussion of benefits should feature so prevalently, or at all, within the field. Whilst a paper discussing benefits regarding political philosophy is not original, a paper aiming at trying to get to the bottom of what a benefit actually is, or at least what it is broadly understood as being, has not yet been done. Thus, this paper will not conclude that there are necessary and/or sufficient conditions which can be derived to decipher what should count as a benefit. Instead, this essay only aims to highlight that there is presently no universal understanding of the term within the literature, yet this fact is often taken for granted. Habitually, as is the case with the notion of benefits, when terms are ubiquitously used they are seldom defined. Subsequently, the problem of there being no widespread understanding of what the term actually means can, and in this instance, does, arise. Clearly, for philosophy to be the most productive and pragmatic, it is imperative that the meanings of concepts are clearly-defined and therefore when they are used in different works, they are being understood in the same way. Otherwise, progress will be stunted and, given that political philosophy is difficult enough, the more ambiguities which can be avoided, the better. Hence, this paper is trying to disambiguate the term and bring to the forefront the many ways in which people can, and indeed do, understand what it means for something to be a benefit. As I will argue, the intuitive and common interpretations of the term are not without their problems, yet I am not concluding that one construal is correct.

Keywords
Metaethics, Ethics, Political Philosophy, Benefits

The Relevance of Disambiguating Benefits
The concept of benefits has long been at the heart of discussion in political philosophy. Many political philosophers, both contemporarily and historically, have
used benefits to ground political obligation and subsequent theories which stem from political obligation. Many philosophers claim that the government having provided us with “benefits”\textsuperscript{1} grounds our political obligation to obey the laws of the state. Benefits are often used as grounding for political obligation which then stems into fair-play theories, consent theories, gratitude theories, associative theories etc. Thus, it is clear that benefits have a key role in the fundamental justification for political obligation and the subsequent theories which arise from it.

It is important at this stage to outline what the aims of my paper are and what they are not. Firstly, my paper is not arguing for any one viewpoint. I have no commitments to whether we should incorporate benefits into our political theorizing. What is clear to me is that benefits are very much a part of political philosophy and that is why they are the subject of my paper. Similarly, I am not arguing for or against political obligation or any theories which incorporate it. However, the link between benefits and political obligation justifications is well-documented in the literature, and therefore requires philosophical investigation independently of whether discussion of benefits should feature so prevalently, or at all, within the field. Whilst a paper discussing benefits regarding political philosophy is not original, a paper aiming at trying to get to the bottom of what a benefit actually is, or at least what it is broadly understood as being, has not yet been done. Thus, this paper will not conclude that there are necessary and/or sufficient conditions which can be derived to decipher what should count as a benefit.

Instead, this essay only aims to highlight that there is presently no universal understanding of the term within the literature, yet this fact is often taken for granted. Habitually, as is the case with the notion of benefits, when terms are ubiquitously used they are seldom defined. Subsequently, the problem of there being no widespread understanding of what the term actually means can, and in this instance, does, arise. Clearly, for philosophy to be the most productive and pragmatic, it is imperative that the meanings of concepts are clearly-defined and therefore when they are used in different works, they are being understood in the same way. Otherwise, progress will be stunted and, given that political philosophy is difficult enough, the more ambiguities which can be avoided, the better. Hence, this paper is trying to disambiguate the term and bring to the forefront the many ways in which people can, and indeed do,

\textsuperscript{1} As this paper discusses, it is apparent that there is no uniform understanding of the term. I have therefore placed it in quotation marks initially to highlight this point, yet for ease of reading will refrain from repeating the quotation marks throughout.
understand what it means for something to be a benefit. As I will argue, the intuitive and common interpretations of the term are not without their problems, yet I am not concluding that one construal is correct.

**Applications of the Term within the Literature**

To elucidate the point that discussion of benefits is customary within political philosophy, this section will present ways in which the term is used in the literature. Socrates is often viewed as the first to defend the view that receiving benefits gives rise to political obligation². “Since you have been born and brought up and educated, could you say that you were not our offspring and slave from the beginning, both you and your ancestors?” (Plato, 2012, 50e). Whilst presenting an extreme position, Socrates is vocalizing the view that Plato having received an education, and indeed perhaps simply being born into the given society, means he must accept the laws. In Plato’s instance, this meant accepting being sentenced to death. Hence, benefits have been at the core of discussion of political obligation from very early-on in philosophy.

More recently, Rawls used the acceptance of benefits to ground his principle of fair play. “A person who has accepted the benefits of the scheme is bound by a duty of fair play to do his part and not to take advantage of the free benefits by not cooperating” (Rawls, 1964, 9). Whilst Rawls is addressing his view regarding the impermissibility of free riding, and agreeing with Hart that “if others are cooperating for mutual benefit and I benefit from their cooperation, then I have an obligation to do my share” (Hart, 1955, 185-6), he is clearly using the concept of benefits to justify his claims. According to M.B.E Smith, fair-play theorists such as Hart and Rawls “conclude that those who benefit from such legal systems have a prima facie obligation to obey the law”³ (Smith, 1973, 955).

Importantly, if benefits ground political obligation and any principles or theories which are derived from it (such as the duty of fair play), this will have a prodigious impact on what people are justified in doing with regard to following and breaking the law. If one is bound in every instance to follow the letter of the law because they have received benefits, this will deny people the right to engage in civil disobedience,

---

² Political obligation can be understood as “to have a moral duty to obey the laws of one’s country or state” (Dagger & Lefkowitz, 2014).

³ A person S has a prima facie obligation to do an act X, if, and only if, there is a moral reason for S to do X which is such that, unless he has a moral reason not to do X at least as strong as his reason to do X, S’s failure to do X is wrong. (Smith, 1973, 951).
even if the government and its laws are entirely unjust and immoral. Even if some circumstances are such that they defeat political obligation, if the notion of benefits play a fundamental role in determining if and when someone must abide by the law, it is of paramount importance that we understand what benefits are. According to Socrates, receiving an education is a benefit which results in political obligation, and perhaps simply being born into a given country is also a benefit. Rawls states that accepting benefits commits us to contributing our fair share, yet is silent on what benefits he has in mind. Is it the same benefits as Socrates held, such as education? If so, do people who have not received a satisfactory education have no obligation to obey the law? What if I had an education that I did not want thrust upon me? The notion of “mutual benefit”, as mentioned by Hart, is a tenebrous notion in-and-of itself, as it appears to depend on a shared understanding of a benefit and for everyone to benefit in the same way; both of which conditions appear entirely unattainable from the outset. What is immediately apparent from considering a few scholars’ uses of the term benefit is that there is no clear way to understand the term. In the subsequent sections, I will undertake a conceptual analysis of the term and try to uncover problematic ways of interpreting it by presenting counterexamples directed at the intuitive ways we understand benefits.

**What Is a Benefit?**

At first, the question “What is a benefit?” may appear obvious and unworthy of scholarly discussion. Yet, immediately, it is very hard to provide a definition which one is comfortable accepting as having encapsulated all of the necessary components of a benefit without also including erroneous or unsettling commitments. Plainly, what it means to benefit, gain, or profit from something is not remotely obvious. To begin, let’s take an initially uncontroversial example of a benefit. Alan gives Anne $100. Anne now has $100 more than she had before, and can now buy the designer jacket she had been eyeing up. Anne certainly seems to have benefited from Alan’s act. Thus, shall we simply conclude that the mere act of getting something you did not have before is a benefit? Clearly, one can be given something that is harmful for them and does not benefit them in the slightest, such as a black eye.

To elucidate this point, let us now imagine that, unbeknown to Alan, Anne has a severe heroin addiction, and the $100 he provided her with went entirely to her

---

4. All of these issues will be discussed in the upcoming sections, yet they are relevant to bring up here also.
substance abuse problem and ultimately led to her premature death. Evidently, Alan giving Anne $100 did not benefit her in the slightest here, yet she still received the same $100 which she had before. Of course, perhaps getting $100 from Alan did benefit Anne in some sense. Anne wanted heroin, and the $100 benefitted her (at least initially) by enabling her to do what she wanted.\(^5\) However, we clearly do not want to say that the act of giving something to another automatically benefits them. Hence, there may be a temporal or causal component to take into consideration. If Alan punches Anne, she is immediately harmed. If Anne goes for a massage, she immediately benefits from the instantaneous relaxation which it provides.

Therefore, we may be inclined to say that if A causes B (either directly or indirectly) in a way which increases B’s wellbeing, A has benefited B. However, this type of definition of a benefit immediately runs into issues of how we will define wellbeing, and whether people can be misguided in terms of what they believe will increase or decrease their wellbeing. Moreover, the causes of events can, in many instances, be numerous and difficult to pinpoint, especially when we begin to trace them back in time. For example, did the owner of Volkswagen benefit me because I own and drive a VW beetle? If we become strict with tracing a causal chain, we appear to be committed to claiming that he did benefit me. Furthermore, this example brings up the problem of whether or not the benefactor must be aware that they are providing a benefit, which we will discuss later. Clearly, the owner of VW has no idea who I am or that I even own one of his cars. Therefore, defining benefits in terms of merely receiving something or by appealing to the notion of causes is problematic even although it initially appears very intuitive.

What the notion of causes and specifically the example of Anne’s heroin addiction brings to the forefront is the notion of consequences. As we have uncovered, simply being handed the $100 is not enough to merit the conclusion that Alan’s act benefitted her. In the instance of the money directly contributing to her overdose, it seems clear that consequences matter. However, what if as soon as she is handed the $100 bill it blows away in the wind? Now, Anne, with or without a drug problem, cannot spend the money. Did she still benefit from simply being handed it, or did she not benefit due to not being able to use it? In my opinion, if Anne cannot do anything with the money, she has not benefitted by having had it for a split second. Indeed, it is perhaps intuitive

\(^{5.}\) I will discuss the issue of wanting something and whether or not that should feature in our definition of a benefit later on in the essay.
to claim that having the money for such a short time then it being taken away harms Anne in some sense as she will be haunted by ideas of what could have been.

Therefore, consequences matter, but how should we understand consequences? We do not want to adopt such a strict notion of causes and consequences that we end up saying Alan caused Anne’s overdose by giving her $100, especially if he had no idea she was an addict. Yet, without Alan and his gift, Anne would never have had the money to get her next fix. Understanding benefits in terms of causes and consequences leads us to uncomfortable statements such as claiming Alan is responsible for Anne’s death. Plainly, nobody can predict the future, and to only consider benefits by appeal to consequences presupposes that people are fortune tellers, or that they somehow should be. Alan has no way of knowing if Anne will spend the $100 on heroin, parents have no way of knowing if private schooling will benefit their children more than public schooling because they do not know the consequences until they have embarked on one or other of the paths. Thus, resorting to causes or consequences will always incorporate an element of luck which, if avoided in our theorizing, will sit more comfortably.

Related to the notion of consequences are intentions. If Alan did not intend for Anne to spend the $100 on heroin, does that matter? Did Alan still benefit Anne in some sense because his intentions were pure, irrespective of consequences or what the $100 caused to occur? As we have discussed, if Alan did not know Anne would use the money to overdose, we probably should not hold him responsible. However, if the money blows away before Anne can buy her designer jacket, Alan’s gift does not appear to have benefited her. Yet, imagine Alan did know that Anne had a drug addiction and hoped she would use the money to overdose. In this case, his intentions are deplorable, as he is hoping his money will contribute to her demise. However, suppose Anne uses the $100 to go towards drug rehabilitation and eventually gets clean. Without Alan’s $100, Anne would never have had the funds to afford the treatment program which saved her life. Did Alan’s gift benefit her? In some sense, it certainly did, as good results came from it. However, if Alan intended to do harm to Anne, it would seem incongruous to claim that he benefitted her in some way. However, perhaps the fact that he gave her the money is irrelevant in this instance, given his intentions. If that is the case, do intentions override causality? Moreover, it would be even more contentious to argue that Anne is indebted to the immoral Alan simply because his intentions did not come to fruition.

In the instance where Alan is hoping for Anne to use the money for drugs, it is perhaps relatively obvious that he has not benefitted her by giving her $100, especially
if she indeed uses the money to fuel her recovery. However, let’s assume that Alan is a multimillionaire who is providing $100 bills to strangers in order to feel good about himself as those who see him will view him favourably. According to Smith, “if someone confers benefits on me without any consideration of whether I want them, and if he does this in order to advance some purpose other than promotion of my particular welfare, I have no obligation to be grateful towards him” (Smith, 953). In this instance, Alan benefitting Anne is only a fortunate by-product of Alan benefitting himself and promoting his own agenda. Does he really benefit Anne if he has no concern whatsoever for her welfare but is solely interested in his own? It seems that in order for someone to benefit another they must at least be concerned with them to some degree, and not in a negative manner. Alan was concerned with Anne’s welfare when he had ill-intentions towards her, yet concern of that nature does not merit benefit. According to Boran, if a good is not “produced intentionally by those who are producing it” then it does not generate obligation. He further writes that “it may seem morally intuitive to hold that intended and unintended benefits are not morally on a par” (Boran, 2006, 105). Yet, this presupposes that one can benefit even if there is a lack of intention from the giver.

Moreover, according to Nozick “one cannot, whatever one’s purposes, just act so as to give benefits...you certainly may not [expect repayment of some kind] do so for benefits whose bestowal costs you nothing” (Nozick, 1977, 166). Interestingly, Nozick claims not only that intentions matter but that there must be some element of cost involved on behalf of the person who is the potential benefactor. In the instance where Alan is a multimillionaire, this fits well with my intuitions. It seems that Alan’s actions are certainly less praiseworthy if it costs him absolutely nothing to give Anne the $100 bill, especially if his intentions are entirely self-interested. However, all that this tells us is that “there is no normative basis” for Alan’s actions and that “we do not usually think that we are under any obligation to reward people for a benefit they produced unintentionally, when doing something they already prefer or value, an activity that is not a burden, but an enjoyable experience” (Boran, 2006, 105). Thus, all that discussion of intentions provides us with is a reaffirmation that unintended benefits where there is no cost involved to the benefactor are not morally on a par with intended benefits which include cost. Unfortunately, it does not aid us in getting any clearer on what a benefit actually is, which is our project.

Moreover, the notions of want and need should not be omitted from our discussion of benefits. If Alan gives Anne $100, whether or not he has good intentions, if Anne is already a multi-millionaire, does she benefit from this extra cash? It would
appear that if she does it is in an extremely limited sense and she may not benefit from it at all. If Anne desperately needs the $100 in order to contribute towards life-saving surgery, she certainly seems to benefit from the additional dollars. Even if we take Alan’s intentions out of the equation and if we omit any resulting consequences from the money from discussion, it seems that whether Anne wants or needs the money does indeed matter. However, the Anne that is addicted to heroin has the want for heroin, and so it would seem wrongheaded to claim she benefits because Alan’s gift satisfies a desire. Moreover, claiming that someone must want something in order to benefit from getting it immediately runs into problems. Very few people want to go on a long-haul flight, sit in a dentist chair or go to the gym every day, yet people certainly appear to benefit from doing these activities. If there was no benefit from going to the hospital and enduring injections, nobody would voluntarily ask to have needles thrust into their arm. One might claim that the benefit is indirect and therefore whilst it is not instantaneous, the benefit is increasing health or seeing family or ridding oneself of pain. This seems true; if one wants to have a painless tooth one endures the pain of the dentist, and one ultimately benefits from sitting in the dentist’s chair. However, Anne the heroin addict, may not remotely want to go to rehab. She might know it will save her life to do so but she has absolutely no desire whatsoever to “get clean”. Regardless, it will certainly benefit her in some sense to get clean, yet it will also benefit her to continue to do what she desires, even if what she desires will harm her. Therefore, there are immediately problems with regarding benefits as needing to be wanted by the receiver.

Clearly, things we want can harm us. Moreover, things we do not want, or things that even harm us, can ultimately benefit us. For example, imagine Sally who is forced to go on a long journey to see her distant relatives in Idaho. She does not remotely want to go for many reasons. She will miss her friend’s wedding, be very behind on her work and whenever she travels she falls victim to terrible travel sickness. In addition to all of this, Sally has no interest in meeting her distant relatives as she knows they are all racists. Clearly, Sally does not want to go to see these family members; in fact, she bitterly objects to her parents making her accompany them. Sally has a horrible time with her racist relatives, vomits multiple times on the journey and is outcast by her group of friends at home for not being at the wedding. There are no perceivable benefits that arise from her trip to Idaho, not even in the months or years that follow the event. We might even be justified in claiming that the trip harmed Sally in some ways. However, twenty years down the line, Sally impresses a man she is on a date with by knowing the capital of Idaho, as it is where her racist family lived. Had she not spent
that awful week in Boise, she would not have known this fact. Did Sally benefit from the trip twenty years ago? It seems difficult to say. What if this man eventually went on to be her husband and knowledge of American State Capitols was a trait he required in potential partners? Whilst we might question the sanity of Sally’s spouse, we may have to concede that she benefitted from her horrific trip to Boise. Yet, this seems very odd to conclude. Perhaps Sally had blocked the awful vacation from her memory and never knew the origins of her Idahoan knowledge. Moreover, if her marriage went on to be unhappy or very pleasant, would that matter? What is clear, though, is that some acts which we may classify even as harmful can eventually give rise to benefits, even if the individual in question is never aware of this fact.

Whilst it is often implicitly accepted that harms are the opposite of benefits, it is important to make this point explicit. Benefits and harms are therefore comparative notions “involving the idea of people being worse off, or better off” (Pogge, 2004, 26). Given that our earlier attempts to understand benefits have been problematic, thinking of benefits in terms of whether one is “better off” to the way they were before they received something seems like a logical way to proceed. The earlier example of Sally brings up the question of whether or not one must be aware that they are receiving benefits. According to Ewing, “the obligation to one’s country is more analogous to the obligation to our parents…the debt is not incurred deliberately” (Ewing, 1947, 218). Often, receiving benefits from the government is held to be very similar to the benefits which one receives from one’s parents. This is in large part due to the fact that there is frequently no realisation that the benefits are being received in both instances. For example, unconditional love does not tend to be regarded as a benefit that comes from parents, but if it has always been that way, it will only seem like a harm if it is taken away. Generally, when things have always been a certain way we do not view their continuation as a benefit. This may be in large part due to us not viewing ourselves as “better off” than we would have been or were in an alternative and prior state.

We certainly speak in terms of being “better off” in everyday conversation; Anne is “better off” for receiving $100 and buying clothes. Yet, thinking of benefits as making us “better off” is immediately problematic. For example, African-Americans are certainly better off living in America today than they would have been living in the country 100 years ago. However, we could not infer from this uncontroversial claim that the current state of racial affairs in the USA is benefitting African-Americans. Similarly, Jewish individuals living in Germany today are certainly better off now than they would have been during the Nazi regime. Yet, this does not entail that they benefit in some
sense from the current ordering of German society. According to Pogge, “drawing this inference [that being better-off entails a benefit], we would beg the whole question” (Pogge, 2004, 26) by assuming the earlier state as the appropriate baseline. Plainly, the fact that racism is falling does not justify the claim that those who currently fall victim to it are benefitting in some sense from this fact. Therefore, it seems that diachronic, or temporal baseline comparisons, do not help us in deriving an unproblematic notion of a benefit.

Additionally, the more historic the baseline for comparison, the more problematic. For example, take the claim that African Americans in America today are no worse off than they would have been had they never had contact with people outside of Africa. Immediately, we should question how much can truly be known about how things would have been in Africa had Africans not been forcibly removed from their homeland so long ago. Such a claim cannot tell us anything very relevant about whether or not African Americans benefit from being in today’s society, as we are merely imaging a way things could have been so long ago. Thus, subjunctive comparisons are equally as problematic as diachronic ones. The more historical our baseline, the more hypothetical our theorising and the less relevant and pragmatic our conclusions become. Theorising about how Africa would have been had there not been colonization and enslavement is as hypothetical as comparing the current state of affairs to an entirely fictional alternative. For example, it would be preposterous to claim that the Nazi regime benefitted people because under the fictional Nazi regime*, one million additional people died. Depending on purely (or even largely) hypothetical examples tells us nothing about the current state of affairs. This is due to the fact that there will always be an X (hypothetical or entirely fictional) where X benefits us more than the current state of affairs Y does. Yet, using X as a baseline for comparison tells us nothing about whether Y benefits us or makes us better off, and it also presupposes that X is the relevant and correct baseline for comparison. Thus, “baseline comparisons do not afford a promising ground” (Pogge, 2004 27) for understanding the notion of a benefit and therefore thinking of benefits as making us “better-off” is problematic, as we are immediately subject to the all-important question “better-off than what?”.

To summarize thus far, benefits cannot be unproblematically understood in terms of merely receiving something or in terms of causes or consequences. The temporal element of causes and results also brings up questions of whether benefits need be immediate, or how postponed an eventual benefit can be. Moreover, it seems that some things which harm us can eventually benefit us. Defining benefits in terms of intentions and costs is also problematic, as not only can we not definitively know.
someone else’s intentions, neither can we predict the outcomes of events. Considering benefits to require the receiver to want the thing in question is also troublesome as we can want things that are bad for us and, as mentioned, things which harm us initially can ultimately benefit us. Furthermore, making comparisons involving a baseline is problematic as the baseline we choose appears unavoidably arbitrary, and there will always be a hypothetical state preferable to the current situation so this fact alone tells us nothing valuable about benefits.

**Summary Conclusion**

At the very least, this paper has uncovered that all of the intuitive ways of understanding the notion of a benefit are not without their problems and counterexamples. It has also shown us that a concept which is so frequently used in both everyday and philosophical discussion is not universally understood. Given the notion of benefits is so confused, perhaps an argument could be made in support of omitting benefits from discussion in political philosophy. However, given that the use of the term is currently ubiquitous, I trust that this essay will highlight to scholars that we cannot assume that our readers or other academics understand the term in the same way. Moreover, being explicitly aware of the problems which our chosen interpretations possess will always make for much more cautious philosophy. Ideally, by scholars who write about benefits being aware that there are many varying interpretations of the term, we might eventually uncover the correct way of understanding it. Perhaps the best and least problematic manner is one which is not intuitive and therefore has not been thought of yet. By assuming scholars view terms in the same way and taking concept’s definitions for granted, we can never be certain that we are viewing notions in the correct, or most beneficial, way.⁶ This essay has shown that there are multiple competing ways of understanding a benefit, yet accepting and receiving benefits plays a fundamental role in grounding political obligation which is a key concept in political philosophy.

**References**


---

⁶. Pun intended.
Stevenson


