Breaking Down Dirty Hands: Episodes, Careers, and Rules

Sarra, Gianni

Awarding institution:
King's College London

Breaking Down Dirty Hands: Episodes, Careers, and Rules

Gianni Sarra

KING'S College LONDON

A Dissertation Submitted in Fulfilment of the Requirements for the

Degree of Doctor of Philosophy

Department of Political Economy

King's College London

June 2023

Word Count: 88,368

Abstract:

This thesis defends and reconstructs the phenomenon of all-things-considered (ATC) justified moral dirt within politics. More specifically, I propose a tripartite typology of the different sources of this dirt. ATC-justified dirt can arise from different pressures and features of our moral universe, and understanding what is behind a kind of dirt being ATC-justified is crucial for properly understanding it: for example, it helps to assess whether future instances of this kind of dirt being ATC-justifiable are preventable. These three categories, in short, are Dirty Episodes (discrete choices where a wrongful act is justified because it is better than the alternatives); Dirty Careers (the moral deviations demanded by a political career even in ordinary and functional times); and Dirty Rules (where the 'terms of the competition' unjustly require or incentivise dirtying behaviour).

My project, in essence, is a reformulation of the Dirty Hands (DH) framework. The DH framework posits that there will be decisions, mostly of a political nature, where an agent must do wrong in order to do right; in other words, that in some situations the best course of action requires committing a tangible wrong. Often DH theories present DH problems as necessitating a certain kind of structure: for example, Michael Walzer presented DH conflicts as involving determinate wrongdoings within a discrete decision, where pressing consequentialist imperatives overrule deontic constraints. I agree with those critics who reject the focus on these factors as necessary conditions of DH.

In this thesis, I instead focus on the underlying concept of justifiable dirt, detached from any specific framework of how it might come about. I argue that this notion can explain much of what is morally troublesome about political life, including issues that the orthodox DH account struggles with. I define dirt as being instances where there is a 1) uncancelled moral ought (expressed in terms of conflicts between plural and conflicting values), that 2) grounds

legitimate criticism of both the act and the agent, and 3) leaves behind a tangible moral remainder. I then analyse the different structures of choice that can leave an agent justifiably accruing this dirt.

In the first two chapters I defend the coherence and theoretical attractiveness of the notion of ATC-justified dirt. By examining conflicts between plural and conflicting values, we can identify the conceptual space for acts that contain both wrongful and rightful elements. Allowing for ATC-justified dirt is an especially important within politics. There are many reasons for this: the moral significance of political goals, the role that the political process plays in determining a society's values, and the need to make effective political participation a permissible option for morally decent people. There is also a need, especially important within politics, to allow deontic constraints such as rights to retain evaluative importance even when they are broken. There are attributes of politics that mean that there will be many ways in which dirtying acts often emerge as necessary or ATC-desirable. These include the role of moral luck, the inescapable dirt found in many specific dilemmas, and the role of democracy itself.

I defend the notion of ATC-justified dirt against accounts that deny its existence. Firstly, I reject purism—an absolutist commitment to certain deontic constraints—as often being undesirable or impossible to achieve. Secondly, I turn to specificationism, which argues that when faced with a seemingly dirtying conflict, we need to change how we view one or both values to dissolve the conflict. While specificationism will often be the right approach, sometimes it is inappropriate: not all conflicts can be redefined in a way that leaves no moral remainder. Thirdly, I turn to threshold deontology and consequentialism. These two views are separate, but they both share a common theme: when it is right to break a deontic norm, no true moral dirt attaches to the agent. I argue that this view is not only incorrect (based on how moral remainders can arise from value conflicts) but deprives us of a useful resource for understanding what is truly at stake in such choices.

While I defend the notion of ATC-justified dirt, I agree with some of the criticisms of the DH framework, or at least criticisms of how it has been used, and build on these to challenge two prevalent positions. One is the orthodox DH view, which focuses on single moments of shocking personal choice that arise from something dysfunctional or disastrous in a political situation, and the other is a dynamic view, where politicians must leave traditional morality behind altogether and enter a Different World, morally speaking. Neither of these accounts, I argue, can explain on their own what is dirtying about politics. They ignore the importance of broader contextual analysis and how political ethics can entail complex mixtures of paradoxes of both character and action. Discussions of DH should not be reduced to either of these views, but nor should either view be excluded from our discussions of DH completely. Both views should have a role, but even together they do not capture all iterations of ATC-justified dirt in politics.

In the final three chapters, I set out my threefold classification of different forms of dirt. Firstly, there are Dirty Episodes (DE). These most resemble the orthodox DH view, covering cases where there are discrete or delineable decisions that represent some dysfunctional break from ordinary politics, in which a determinate wrongdoing can be justified. While these cases do not cover the entirety of dirt within politics and too often have been discussed in an unduly static and individualistic way, they remain important. Key examples include swift decisions made in the context of an eliminative emergency and how agents respond to the corrupt offers put forward by others.

Secondly, Dirty Careers (DC). This is the familiar notion that, in seriously pursuing the profession of politics, one has to accept the self-justified norms and dispositions of politics, and that this process can be dirtying. Examples include the demands of partisan affiliation and political rhetoric. Finally, I propose the novel category of Dirty Rules (DR). This covers instances where the reason why agents feel they must get their hands dirty is institutional. There

is something remedially and contingently wrong about a political climate or ruleset; a structural failure that is dirtying for those who engage with it. Examples include redistricting systems that incentivise gerrymandering, campaign finance systems that require corrupt dealing, and polarised political climates that legitimise more aggressive tactics.

This threefold classification allows cases of ATC-justified dirt to be analysed with greater accuracy. It helps to identify how immutable and inevitable a given moral conflict may be for political actors, and also whether or not they represent something dysfunctional within a political climate.

Acknowledgements:

Writing this thesis would not have been possible without the help and support I received while creating it. Listing everyone who helped, in ways big and small, would be an impossible task and so I can only apologise to all those who I will inevitably neglect to list within this section. Lack of individual mention does not mean that I am any less grateful for what they have done.

It is only right to start by thanking my supervisors, Robin Douglass and Adam Tebble, for guiding me through this process and providing me with encouraging words of wisdom about both this project and my ambitions in academic political theory more broadly. In particular, I want to thank Robin for the help he provided in the application process and the many insightful and thorough comments he provided on various drafts of my thesis. The speed and diligence with which he provided feedback always consistently surprised me and was essential in getting this project to this stage.

Long before I even began work on this thesis, my interest in political theory was sparked as an undergraduate student at this department almost a decade ago, and both Robin and Adam were a part of that. I look both fondly on the undergraduate political theory modules that were convened by Adam and the experience of writing an undergraduate dissertation with Robin as my supervisor.

The Department of Political Economy proved an invaluable source of support and assistance, and I owe thanks to many figures throughout the academic and professional services staff. To single some people out, I want to thank Yuchen Sun and Dulyaphab Chaturongkul, two of my fellow PhD students in political theory with whom I have had many thought-provoking conversations; Steven Klein, Billy Christmas, and Marta Wojciechowska, for providing me with invaluable experience in teaching and marking; the GTAs who I worked alongside; Carmen Pavel for helping make the upgrade a rewarding experience; and the Political Theory Workshop more generally for providing crucial experiences and feedback.

Table of Contents

INTRODUCTION

Few professions invite moral criticism and contention to quite the same degree as politics. This is perhaps unsurprising. Politics involves questions of justice and rights, and momentous decisions that affect the lives of countless others. If morality is needed to circumscribe the boundaries of any sphere of life, the realm of politics should be a prime candidate.

However, in reality a moralistic attitude to politics seems to be a recipe for disappointment, irrelevance, or failure. Successful politicians are rarely lauded, and still more rarely *correctly* lauded, as exemplars of moral propriety. Politics is a dirty craft that requires dirty behaviour. Philosophy, art, literature, popular culture, public discourse, and so on are filled with descriptions of the sly and Machiavellian politician, climbing the greasy pole in a dog-eat-dog world. Politicians rise and fall based not on ethical standards or moral deservedness, but on their ability to play this sordid game.

We may disagree about *why* politics has this dirty moral reputation: is it a sad empirical reality, the contingent result of how our societies have been organised? Is it an inevitable part of politics, whether because of inescapable political realities or because the nature of politics itself generates its own normative standards distinct from conventional morality? Could we conceive of a better politics that frees politicians from this morally compromised world? Or is this the wrong question altogether, and might there actually be something desirable about having politicians who are willing to get their hands dirty? In short, we disagree about whether dirt in politics is inevitable and desirable. Few would argue that the self-interestedly corrupt politician or the aspiring tyrant are justified in their actions. Some would argue, however, that the so-called heroes of politics can permissibly, or are even morally obligated to, use methods that resemble those used by the villains.

In this thesis I side with those who think that politics can create situations where a political actor is justified in doing something morally wrongful, in situations ranging from the mundane to the exceptional. This view is associated with the "Dirty Hands" framework, where the best course of action or the right way to relate to one's role can create tangible and meaningful moral remainders. The option that is all-things-considered (or ATC) justified might still be, in an important way, morally wrong. Although we do not want to be ruled by politicians who get their hands dirty *for the wrong reasons*, nor do we want to be ruled by ones who value their personal moral cleanliness above all other priorities.

My main argument in this thesis is that there are different kinds of justified moral wrongdoing and that the *justified* dirt attached to politicians takes different forms, differentiated based on what renders this dirt ATC-justified. The dirt generated from the justifiable immoralities within politics, whether understood as resulting from paradoxes of action or paradoxes of character, has too often been treated as a monolithic concept. In reality, I argue, there are different ways in which this dirt can become ATC-justified, and there is no single structure of choice that captures all DH decisions.

A traditional way of breaking down dirt has been based on the severity of the dirt inflicted: how serious was the moral norm that was broken? I go beyond this by also breaking dirt down based on its source and by identifying what renders a dirtying act ATC-justified in the first place. By breaking down this monolith, I also bolster the 'dirt in politics' thesis against common criticisms and help to identify its distinctive role in political ethics.

To do this, I introduce a tripartite typology of the different forms of dirt: Dirty Episodes, Dirty Careers, and (most novelly) Dirty Rules. The 'dirty hands' literature falls into the all-too-common trap of talking about the 'necessary' immoralities of politics as if they are one and the same. The literature usually focuses on what I call "Dirty Episodes", where a specific decision requires doing something that would normally be impermissible in the name of preventing a

14

lesser evil. However, what about those scenarios where the quotidian daily functioning of politics requires the mundane cultivation of morally wrong practices and techniques? Politics can be described as a "Dirty Career" (hereafter DC). The morally distasteful no longer becomes justified due to aberrational events, but because of ordinary day-to-day necessities. Though this idea has seen some discussion in political theory as a more dynamic alternative to the orthodox Dirty Hands framework, I emphasise that while it requires distinct treatment from Dirty Episodes (hereafter DE), they are not mutually irreconcilable concepts.

One specific type of moral conflict that I think the literature is currently ill-equipped to deal with are what I call "Dirty Rules" conflicts. These arise when a specific institutional environment means that immoral behaviour is legitimised, incentivised, or even required. From the standpoint of institutional design, Dirty Rules (or DR) systems are bad, and creating or entrenching them is wrongful. There is a moral imperative to change these systems and a moral wrong involved in leaving them in place. They are, after all, remediable. They are structural contingencies that agents must currently operate within, but they could be changed. Many political actors, pursuing both permissible and morally important political projects alike, will find themselves operating within DR systems. Effective political action, including reform efforts, requires playing by the rules as they are, not as how they should be.

DR conflicts arise when playing by those rules entails morally dirtying acts where the dirt falls, in part, on the agent *qua* individual. This dirt goes above and beyond what is inflicted by the 'career' of politics. How best to achieve political success, perhaps even what political success means, is significantly more morally compromising than it needs to be. A morally clear-headed politician will recognise not only that the system is unnecessarily flawed, but that it asks aspiring politicians to become flawed as well. The agent cannot, under these conditions, place the moral responsibility entirely at the foot of the institutional environment or its authors.

What is the just politician to do? Should they accept, perhaps even internalise, the rules of the game as they currently exist until they have the power to change them? Or should they stick to their principles, morally rising above the system? Different politicians will answer this question in different ways, but it is generally best to have politicians willing to accept and work within these constraints. Doing so is, however, dirtying. A moral residue is created; an uncancelled immorality remains, even if permissible or justified. Despite this shared dirt, DR practices should be treated differently both to DE and DC acts: DR scenarios, like DC and unlike DE, represent the rules of the game, but should still not be normalised.

1. The Dirty Hands Debate

Dirty Hands (hereafter DH) dilemmas represent more than ordinary moral conflicts or the standard trade-offs between important values. They refer to when an act is both right to do and still meaningfully wrong (Walzer, 2004: 33). DH acts contain both rightful and wrongful elements that do not wash each over away (Goodwin, 2009: 16-17; 144-147). DH dilemmas arise in "those unfortunate, difficult, and sometimes tragic moral situations, where no matter how the agent[...]acts there can be no avoiding the violation of persons and/or violation of a cherished principle" (De Wijze, 2018: 132). They often involve acts that would otherwise be evil (Kramer, 2018: 187), such as enabling the "evil projects" of others (De Wijze, 2007: 11-16).[1] An act that "[violates] an important moral constraint" (Roadevin, 2019: 123) becomes justifiable or perhaps even obligatory, since the costs of *not* violating these norms become too great to bear. It becomes right to contemplate an action which "breaches a moral constraint

---

[1] At a more fundamental level, the term has been used to describe many different views which "involve the idea that correct political action must sometimes conflict with profound moral norms" (Coady, 2018). The most enduring and conceptually distinct definition within the literature, however, focuses on these stark conflicts and lamentable situations wherein a politician is *morally* justified in breaking a serious *moral* norm. That this conflict between "correct political action" and "profound moral norms" can be more quotidian is another recurring view, one that is best understood as a distinct yet related DC claim.

thought normally to be absolute", and the character of this wrongdoing is "not expunged or silenced" by these circumstances (Archard, 2013: 778). Usually, the justification for such acts falls outside what will normally be permitted by the cost-benefit analyses of political decision making.

Key to this is the idea of a moral remainder. There is an "uncancelled moral disagreeableness" (Williams, 1978: 62-63), a "residual moral claim" against the agent (Hall, 2022: 218), one not attended to by the decision's overall justifiability. The agent does not completely "get off the moral hook" (Williams, 1973a: 183) with regards to the specific moral obligation that they broke. In most moral dilemmas, moral remainders usually arise because of the dilemmatic nature of these choices, where there is no 'right' answer. What sets DH choices apart is that the dirt attaches even to the unambiguously right answer (Coady, 2018: 5). Even the right thing to do creates a real moral blemish, tainting the agent's moral character and grounding legitimate moral complaints, which is not washed away by the overall justifiability of the act.

There are a multitude of ways in which this dirt can be described: DH agents compromise their personal integrity (Mendus, 2009: 29-32) and the integrity of broader society's politics (Dovi, 2005: 132); dangerously jeopardise the rules that shape the broader social order (Klockars, 1980: 36); shed any claim to innocence (Hampshire, 1989); commit wrongs against others (Kamm, 2000) in a way that can enable legitimate resistance and retaliation (Smith, 2018: 210-211); heartlessly break the ordinary norms of permissibility (Kramer, 2018: 195); violate deontological prohibitions that apply regardless of appearance or outcome (Bellamy, 2010: 425); break natural virtues of beneficence and compassion (Finlay, 2011: 424-425); "intentionally [further] the aims of evil" (De Wijze, 2007: 16); engage in categorical wrongs such as rights violations (Levy, 2007: 39); frustrate or sidestep proper democratic functionings (Shugarman, 2000; Ramsay, 2000: 36); and commit "morally

repugnant" acts that moral persons will normally recoil from (Curzer, 2006: 34).[2] The dirt in DH conflicts can take all of these forms and many others besides.

1.1 Walzer's Dirty Hands

The modern DH debate originates with Michael Walzer's 1973 paper "Political Action: The Problem of Dirty Hands". Walzer (1973: 160) argued that politicians would often be forced to "choose between two courses of action" where the ATC-right action entails dirt. Politicians must refuse a purist absolutism, but still acknowledge that the dirt they incur exists and is not annulled by consequentialist or political necessity. Citizens, too, want politicians who *within certain contexts and for the right reasons* break these moral norms (Walzer, 1973: 162). DH reasoning provides both excuses and justifications for the moral violation that the politician commits, containing within it both "an admission of fault" and an "assertion" that the correct action was taken (Walzer, 1973: 170).

Walzer (1973: 161) derives the term 'Dirty Hands' from the Sartre (1949) play of the same name, and more specifically from the philosophy espoused by the character of Hoederer. Hoederer, a senior Communist revolutionary, stands accused of trading his ideals for political exigency, acting in manners contrary to his ideology and collaborating deceitfully with class enemies in the name of consolidating political power. This is not an accusation he denies. Instead, he defiantly tells Hugo, a more idealistic party member, that "I have dirty hands right up to the elbows. I've plunged them in filth and blood. Do you think you can govern innocently?" (Sartre, 1949: 244, cited by Walzer, 1973: 171). We must, Walzer asks us to

---

[2] I will discuss these views throughout the thesis. It is worth saying now that I believe, while they all represent true consequences of dirt, their validity varies between the different categories of dirt. For example, a focus on personal integrity is more apposite when discussing DC, while a focus on violating norms of permissibility and stepping outside democratic functionings is more suited for DE.

conclude, accept that the right thing to do in politics is determined not by moralistic principles alone. Politicians must, if they want to achieve good things and avert catastrophes, adapt to how others have "set the terms of the competition" (Walzer, 1973: 163). However, we also do not want to be ruled by politicians who are subject only to "prudential control" (Walzer, 1973: 179). Politicians must not hide their dirt, but instead submit themselves to social control and offer a moral account of their actions. The way to do this, then, is to understand these choices as discrete DH decisions, where a "determinate crime" is committed for which a "determinate penalty" must be paid (Walzer, 1973: 178).

1.2 Differing Views on Dirty Hands

There are some important debates *within* the DH tradition. First is the question of what DH conflicts are actually between. What kind of situation, in other words, can give rise to the possibility of acceptable moral wrongdoing and what does the resulting choice look like? Answers to this vary throughout the literature (Nick, 2019a: 2-4). Often this conflict is presented as the result of competing moral frameworks, both of which are owed some fealty by responsible moral agents and, in DH cases, pull in opposite directions in terms of action-guiding content. "One powerful strand in morality comes into conflict with another" (Coady, 2018: 4) under this account. The resulting dirt therefore arises when the framework not acted upon gives rise to a serious moral remainder.

There are two prevalent competing frameworks views. The first is to view DH as arising when there are dirtying conflicts between consequentialist and deontological moral demands. When deontic norms are broken under consequentialist pressure, dirt is inflicted. This view is favoured by Walzer (1973: 161), and is often treated as the definition of DH, including by critics (for example, Ramsay, 2000: 30; and Nielsen, 2007: 30). A more limited subset of this

view is the "supreme emergency" view (Walzer, 1977), where we are usually governed by certain deontic constraints but these can be waived, with moral remainder, *in extremis* and instead replaced by the "utilitarianism of extremity" (Walzer, 2004: 40).

The other competing framework view presents DH conflicts as resulting from clashes between public and private morality. Individuals can be required to do one thing in their capacity as a private moral agent, but quite another in their capacity as a public officeholder (Nagel, 1993: 79-84). What can be right from the standpoint of public morality can be dirtying in terms of private morality: "public purposes" lead to the violation of still-extant "moral principles" (Thompson, 1987: 11). A similar approach is to argue that, instead of public morality, some idea of necessity, political or otherwise, creates the incentive to justifiably perform dirtying acts: DH conflicts are, under this view, those cases "when morality clashes with some other rational necessity of a profound kind that correctly overrules it" (Coady, 2018: 3), such as political expediency or security.[3] A "severe moral blemish" can thus be incurred in the name of "a politically required course of action" (Meisels, 2008: 154).

An alternative way of viewing DH conflicts is to say that moral conflicts represent clashes between plural and conflicting values (Stocker, 1990; Williams, 1973a). There is no clash between competing moral frameworks under this kind of account. Instead, there are conflicts between morally important oughts and values; moral demands that can exist within the same moral framework yet still be mutually exclusive. This kind of conflict becomes dirtying when a particular threshold of moral severity is passed. For instance, this threshold may be passed when people's statuses as ends are violated in distinct or particularly shocking ways (Stocker, 1990: 17).

---

[3] I do not focus on this view in my thesis, in large part because I believe that these pressures are morally important and should not be viewed as external to morality. This is largely owing to my adoption of a value-conflict view as opposed to a competing-frameworks account that is much more exclusive about what counts as morality (see Chapter 1, Section 2.1). I do, however, want to make some allowance to the idea of non-moral considerations overruling moral ones: this role is limited to instances of DR and DC, where such overruling can be permissible, however, as opposed to DE.

There are other internal debates within the DH literature. For example, there is the question of how avoidable dirt is. Political theorists tend to view these DH choices as either symmetrical or asymmetrical (Nick, 2019a); that is, does the agent meaningfully 'keep their hands clean' by choosing one option over another? If there is a deontic constraint that needs to be violated for consequentialist benefit, does the agent only *truly* dirty their hands if they violate the deontic constraint? Furthermore, DH are often presented as manifestations of poor moral luck. In many cases, that an agent is faced with a situation in which the ATC-justified right thing to do is dirtying is beyond the agent's control. However, if moral luck is true, then it is still right to treat an agent "as an object of moral judgment" (Nelkin, 2023: 1). This is especially apposite if we adopt a symmetrical view: under symmetrical views, there is no way in which an agent can escape with clean hands, and this is a strong case of bad moral luck.

1.3 Critiques of Dirty Hands

The DH framework is not without its critics. Much of this thesis will be spent answering criticisms and restructuring the notion of ATC-justified dirt in response. There are three broad categories of criticism. First, some critics reject the validity of the *project* of DH, questioning the notion of permissible wrongdoing and denying that the right action can genuinely create true moral remainders. Second, some critics take aim at the framework in which DH choices are usually presented, arguing that these theories pay inadequate attention to context or obscure the true nature of political dirt. Third, some critics reject the view that DH can coherently arise within democracies.

Regarding the first category, there are several ways to argue that the core theoretical tenets of DH are incoherent or undesirable. One version of this is purism, also known as absolutism (though I favour 'purism' here as other views could also be interpreted and enforced

in absolutist ways). Strong purists believe that no allowance should be made to the gritty realities of politics, and we should act as moral agents regardless of circumstance or expediency. The weaker form of purism focuses instead on certain deontic constraints that cannot be violated. Another approach is specificationism (Thomson, 1976). Under this view, when two obligations clash, one or both obligations must be reinterpreted to make the conflict go away.

The most interesting forms of this first category, however, are to be found in threshold deontology and (most influentially) consequentialism. Threshold deontologists argue that we should respect deontic constraints up until a point of severe enough emergency where these rules can then be waived. Consequentialists argue that the best action is to be determined based on its consequences, whether this is from maximising utility or minimising evil. What unites these views — purism, specificationism, threshold deontology, and consequentialism — is a rejection of the core idea of DH: that something can be both the right thing to do and simultaneously have a meaningfully wrongful element.

The second category of criticisms focuses on the framework of DH, accepting (the possibility of) permissible wrongdoing yet arguing that the model used by theorists such as Walzer is misleading. One such criticism is that it is unhelpful and misleading to focus on DH *dilemmas* (Tillyris, 2015; 2019a). An episodic focus on single instances of paradoxical choice or tragic conflict, where determinate wrongdoings merit determinate penalties, obscures what is truly dirtying about politics. Instead, the dirt in politics needs to be understood more holistically, as a dynamic tension between lifestyles. The DH framework's focus on "the politics of the great gesture" (Shklar, 1984: 243) needs to be abandoned. There are, too, other ways in which the DH framework can be criticised as decontextualised, treating the terms by which DH dilemmas arise as immutable or neglecting to study the underlying background conditions entirely (Coady, 2008: 89; McDonald, 2000: 188-189).

The third bundle of criticisms concerns the charge that DH acts cannot arise within democracies. Some theorists reject the possibility of Democratic Dirty Hands (DDH). Democratic politics supposedly creates a new political order of the kind that precludes such unseemly acts or prevents individual 'lone wolf' actors acting outside the scope of accountable democratic politics. It is also worth pointing out two theoretical insights that sit explicitly in contrast with this view and suggest that it is wrong to view DH as entailing lone-wolf actors. Firstly, there is the claim that DH acts can be democratically authorised, and as a result the broader citizenry can share in the DH of their elected officials. Secondly, in complex institutions, there will be "many hands" at play (Thompson, 1980) and identifying a single culpable actor may prove difficult if not impossible.

2. Five Examples

To analyse political dirt, and to explore the lacunae in the current dirty hands-dominated framework, five examples of dirtying yet justified moral transgressions will be used throughout this thesis. Thus far, there has been a tendency within the literature to refer to all such examples using the same broad framework, usually a DE approach or a DC analysis. While these five examples all involve dirt, I argue that it would be wrong to treat them as involving dirt *in the same way*. Though there has been work before on distinguishing between different *severities* of dirt, I argue that these examples should primarily be distinguished on another front: what is the reason *why* this wrongdoing seems to be permissible in the first place? This focus on the source of the dilemma is important for determining whether or not such acts should be viewed as anomalous. These examples, it is worth noting, all arise within the context of liberal democracies: how such dilemmas manifest within the context of liberal democracies will be the focus of my analysis here.

Example 1. Dana and the Ticking Bomb Scenario

One prominent DH example that is common within wider discussions of political ethics is the Ticking Bomb Scenario (hereafter TBS; on which see Walzer, 1973: 166-167). In this scenario, a newly elected leader (let us call her Dana), a just and honest politician who sincerely and unequivocally condemned torture, is faced with a terrorist plot. The army she now commands has captured a terrorist who has planted several bombs. The only way to uncover information about the bomb locations, thereby saving innumerable lives by preventing their imminent detonation, is to authorise torturing the terrorist. Assuming that "this is a genuine dirty hands decision and not one flawed by lack of knowledge or ignoring of preferable alternatives" (Coady, 2018: 6), the choice facing Dana is to authorise torture and dirty her hands by violating a deeply sacrosanct universal moral constraint, or to allow the preventable death of potentially hundreds of innocents. Although this thought experiment is very controversial,[4] Walzer's articulation of the TBS is nonetheless useful for bringing out moral intuitions about the options in difficult cases and has received much attention within the DH literature.

Example 2. Ned and the Corrupt Ward Boss

Dirty hands need not be a case of life or death. Another of Walzer's famous examples is that of a political candidate, call him Ned (following Tillyris, 2015: 64), who is greeted with a fundamentally corrupt offer that he must accept to win a morally important election. The deal, to give a dishonest ward boss control over school construction contracts (Walzer, 1973: 165-166), is intrinsically and inescapably corrupt and corrupting. Yet Ned should take it provided

---

[4] As will be seen throughout this thesis. A common source of critique (for example, Shue, 1978; 2006) is to attack the relevance of any moral conclusions that can be drawn from such a strained example.

a certain level of moral significance is attached to his victory. Him prevailing in the election is, in this example, an important enough outcome to justify the wrong that he commits.

Still, Ned frets about his decision for reasons beyond feelings of instinctive discomfort. It violates the trust of those who back him; he worries about his own motives becoming corrupted; he fears the future bargains it might legitimise; he disdains the intrinsic immorality of the act. Those very doubts, those very scruples, give us reason to want Ned – rather than another less hesitant candidate – to make the deal. Those doubts reveal him to be a good man, and we hope that he is pragmatic enough to act against his instincts in the name of the greater good.

Example 3. Edna and the Lobbyist

Edna, like Ned, is a good politician, with all the right scruples and attitudes. To have a viable chance of winning a morally significant election, she is asked to make an unseemly quid pro quo. An industry lobbyist approaches Edna and offers her considerable campaign finance support in exchange for her helping the industry to procure government contracts. On the face of it, this appears to have the same structure as the dilemma facing Ned. A morally undesirable actor is offering a good person an immoral trade, yet this specific person should both take the trade and acknowledge the associated wrongdoing. Their election is substantially more important than preventing the trade.

Unlike with Ned's case, however, what Edna is being offered is entirely legal and normalised under the rules of her political system. The lobbyist is not violating any law or widely agreed-upon professional code of ethics, nor is their act necessarily a secret one immune from public view. Edna advocates the reform of electoral rules to prohibit these trades. In contrast her opponents use this system with glee and without restraint. They do not care for the system's immorality.

The wrongfulness here is a procedural one, with the proper means for resolving such issues and settling democratic disputes being sidestepped or corrupted. The substantive morality, or lack thereof, of the industry in question is not a necessary condition for Edna to feel tormented. The substantive morality of supporting the industry is not the point. Much the same can be said about Ned's example. Like Ned, Edna fears that she might compromise her own character and contribute to broader corruption. Still, she takes the deal and promptly wins the election. Her hands are dirty but, following Walzer, her supporters should welcome this. Though cases such as these are usually described in terms of how to *acquire* power, similar dilemmas to Ned's and Edna's can arise in relation to the *exercise* of power too, calling upon lobbyists and obscure legislative tricks (for Edna) and corrupt extra-legal deals (for Ned) to enact important legislation or defeat unjust proposals. Again, the wrongfulness here is in large part procedural.

Example 4. Paul and the Campaign

Our fourth hypothetical protagonist is Paul. Paul is, by all accounts, a fundamentally decent and admirable person. He is honest, treats everyone as they deserve to be treated, appropriately values important relationships, and has neither the inclination nor the necessary guile to cheat. It is not merely the case that he resists temptations to behave in an immoral manner. Many of these temptations do not even arise, and he is innocent enough not to notice most opportunities for gaining an improper advantage.

Paul loses the first election he runs in. It is a morally important election, and his loss is rightly mourned by his supporters. Upon meeting the realities of a healthy competitive democracy, his strengths become his weaknesses. He loses because of, not despite, his virtues. His commitment to treating his opponents in a fair way renders him an ineffective adversary. His respect for the demands of impartial justice mean that he neglects both the most politically

salient issues in the eyes of the voters and neglects non-morally relevant attributes such as fundraising prowess in what alliances he chooses to build. He is unable to work within the dominant party culture, unable to endorse the party's slate of candidates when he believes better candidates exist in individual elections.

If by chance Paul still managed to get elected, he would find his "virtues" even more debilitating. Not only will he have to compromise on good ideas to get things done, but the success of his political projects will likely depend on unsavoury techniques such as adopting a rather diplomatic attitude towards the truth (for example, blaming opposing politicians for bad news outside their control). His existing disposition is ill-suited for the role which he seeks. There are morally important ends to be achieved using that role, and important moral consequences if his political opponents win. Crucially, he does not think the system is flawed, or at least not in the sense that it *must* be changed. It is a healthy, competitive democracy. Yes, it requires displaying some disrespect towards one's opponents, but nothing dehumanising or abhorrent.

Paul must adopt a different disposition if he wants to be successful. He must learn how to spot and utilise opportunities that he was once squeamish about. This is not a moral 'blank cheque', but he must depart from his conventional views of virtue. In doing so, he feels that he is giving something up in this 'retraining' process, becoming tainted and polluted in some way. If he hopes to succeed, however, he must not agonise about each decision that constitutes this new lifestyle. Furthermore, success is not guaranteed. It might be that he adopts this new disposition, sacrificing his previous virtues, only to fail in politics or belatedly realise he was fighting for the wrong cause.

Example 5. Gerrymandering North Takoma

Gerrymandering is when electoral boundaries are drawn to advantage or disadvantage a particular group, party, or politician. Let us assume we are the leaders of a state branch of Party A, a party in the United States that has recently adopted an unequivocal anti-gerrymandering stance. Decisions on redistricting are devolved to the states. As a result, the constitutional structure permits tactics such as gerrymandering – perverse incentives can conspire to encourage it, even. Party A has recently gained control of North Takoma, a politically competitive state with a high population and an accordingly important number of congressional seats. This is our state, and we now face the decision of what to do with North Takoma's redistricting process.[5]

Our first instinct is to implement a non-partisan redistricting process, ensuring gerrymandering is no longer a political practice within North Takoma. However, something happens which gives us pause. The opposing Party B has recently taken over the state legislature of South Takoma, a neighbouring state with similar characteristics and the same level of political importance. They have mercilessly gerrymandered that state, all but ensuring that, despite South Takoma's overall political competitiveness, Party B will secure a vast majority of the seats that South Takoma sends to the federal legislature and that their control over state politics is guaranteed for some time.

Two factions soon emerge within Party A. One says we must stick to our principles. We cannot fight South Takoma's gerrymandering with a gerrymander of our own. It would be intrinsically wrong to do so and cause tangible damage to democracy within North Takoma. The other faction argues that we must make a limited retaliatory gerrymander. They argue that we are right to accept the discrete and serious wrongs involved in the act of gerrymandering. However, if we do not gerrymander in response to the South Takoman gerrymandering, there are serious consequences for our party's performance in federal elections. Though we

---

[5] I have discussed this example in further detail elsewhere (Sarra, 2022).

acknowledge that it would be fairer overall if there is no gerrymandering, all we can do now is to ensure an approximately level playing field between our parties by mitigating the wrongdoing inflicted by Party B. The only way to do this is by cancelling out the South Takoman gerrymander with a North Takoman one.

There are analogous examples that follow a similar structure, say if morally dubious misinformation practices are legal. To ban them requires getting into power. To get into power requires engaging with the very practices that the agent condemns.

3. The Structure of This Thesis

Chapter 1 investigates what I take to be the bedrock of the DH framework and an essential ingredient in understanding how we think about political ethics: the notion of all-things-considered (ATC) justified moral dirt.[6] This notion is at the core of the three categories of justified political dirt that I describe in this thesis: Dirty Episodes, Dirty Careers, and Dirty Rules. Central to all these concepts is the idea that a political actor can *justifiably* become morally tainted in some way. Many theorists argue, intuitively, that this is contradictory. How can something that is permissible be morally tainting? How can something be simultaneously right and wrong?

Chapter 1 elucidates the idea of ATC-justified moral dirt. I do this first by arguing for the reality of moral conflicts that can, even when the right decision is chosen, inflict moral dirt. Viewing these conflicts as representing clashes between plural and conflicting values easily creates the space for ATC-justified dirt. This is because choosing between conflicting values can create "double-counted impossible oughts" (Stocker, 1990: 12), where we cannot "get off

---

[6] This contrasts with analysis that starts from the structure of the dilemma itself as the core idea.

the moral hook" (Williams, 1973a: 183) with regards to one moral obligation by fulfilling another.

I argue that something becomes dirtying (not *necessarily* ATC-justified dirtying) if three conditions are met. Firstly, an important uncancelled moral ought must not be acted upon. Secondly, the failure to meet this ought grounds legitimate criticism of both the act and the responsible agent, for which they can be held accountable. Thirdly, a moral remainder is generated. If these three conditions are met, then an act becomes dirtying. Given the nature of moral luck and the 'symmetrical' nature of many choices in politics where no option is dirt-free, this is a common phenomenon within politics.

With the notion of ATC-justified moral dirt set out and established as a coherent concept, Chapter 2 defends this idea against its critics, with a focus on how these dilemmas can arise and be defended within the context of liberal democracies. Firstly, I consider the purist objection: that certain moral rules or virtues should always be adhered to, without any concession to political necessity. Given the moral importance of political success and the inability to completely guard against moral conflicts, I reject this view. Secondly, I consider specificationism, which holds that when values and moral imperatives appear to clash, the appropriate response is to reconfigure one or both to dissolve the conflict. While this is sometimes the right approach, I argue against always adopting it. Finally, I turn to the challenges of consequentialism and threshold deontology. These approaches argue that when something is ATC-justified, it cannot inflict dirt. My primary response to these challenges is that it is both correct and helpful to say that dirt can result from even the justified violation of certain rules and standards.

With the notion of ATC-justified dirt defended, in Chapter 3 I propose breaking the concept of DH down further to fully capture the various ethical tensions and quandaries that arise within politics. There is no single appropriate structure for understanding the ways in

which agents may accrue ATC-justified dirt. One relevant distinction already common within the literature is to distinguish dirt based on its severity. I agree with this approach and argue that dirtying acts need to be situated on a continuum from the merely grimy to the bloody. However, we also need to distinguish between incidences of ATC-justified dirt on another basis: what renders such dirt ATC-justifiable in the first place? Why is there a DH dilemma in the first place? What is it, in other words, the source of the DH dilemma? To address these questions, I propose that we distinguish between Dirty Episodes, Dirty Rules, and Dirty Careers.

In Chapter 4, I discuss Dirty Episodes (or DE), which is my account of how the orthodox DH view should be presented. The traditional DH approach has come under criticism for having an undesirably 'static' conception of immorality within politics (Tillyris, 2015: 63), focusing on paradoxes of action to the neglect of paradoxes of character. I agree with this critique and my response is twofold. Firstly, there are important instances of ATC-justified dirt that cannot be understood in an episodic manner. This is why I propose the categories of Dirty Rules and Dirty Careers. Secondly, I argue that we can and should demarcate a space for paradoxes of action. I set out to construct DE in such a way as to do the work that the orthodox DH view is most qualified to do while avoiding the worst excesses of staticity.

In Chapter 5, I examine the dirt that attaches to politics as a career, even in a well-functioning democracy or a just political system. The Dirty Careers (or DC) view covers the diurnal and the quotidian: accepting the need to pursue and value the self-justified goods of politics, changing one's habits and dispositions and committing everyday moral transgressions and hypocrisies without too much pause or hesitation. Even well-functioning political environments can require this kind of mundane dirtying, but I argue that this is not enough to describe all instances of political dirt. While the fact that politics is a DC will be an omnipresent

background condition, there will still be many cases of ATC-justified dirt that cannot be captured by DC.

In Chapter 6, I explore my most novel category of dirt: Dirty Rules (or DR). These conflicts are not the result of episodic disruptions nor are they the costs that attach to even a well-ordered and decent politics. In these cases, dirt arises from contingent structural flaws that there is a moral imperative to fix and a moral wrongness in leaving untouched. Not only should these structures and institutions be changed, but there is dirt that comes from engaging in them. The politician seeking success, however, must engage. I argue that this category has been overlooked by the existing dirty hands literature, which offers no resources for dealing with these DR questions beyond either a DE or DC framework. We need a bespoke framework for such cases. As well as rejecting the notion that DE and DC analyses must exist in an adversarial relationship, I also argue that they are not exhaustive. Viewing certain dilemmas as stemming from DR clarifies both their causes and what is morally at stake.

CHAPTER I: WHAT IS MORAL DIRT?

Politics is a dirty business. This refrain is the starting point for many political theorists in some form or another. Michael Walzer (1973: 162) began his analysis of DH in politics by noting the "conventional wisdom to the effect that politicians are a good deal worse, morally worse, than the rest of us". Hannah Arendt (1967 [2006]: 223), highlighting a specific virtue eschewed by politicians, describes how "no one has ever doubted that truth and politics are on rather bad terms with each other, and no one, as far as I know, has ever counted truthfulness among the political virtues." Judith Shklar (1984: 243), similarly, contends that hypocrisy and betrayals are inevitable and necessary within politics. Bernard Williams (1978: 62) writes that "it is a predictable and probable hazard of public life that there will be these situations in which something morally disagreeable is clearly required. To refuse on moral grounds ever to do anything of that sort is more than likely to mean that one cannot seriously pursue even the moral ends of politics."

Many theorists have argued that there is an abrupt divide between the demands of politics and the demands of conventional morality. For Machiavelli (1532 [2008]: ch. XV), the politician must "learn how not to be good". For Max Weber, politics requires "morally dubious" or "dangerous" methods (Weber, 1919 [1946]: 121). Even those with more moderate positions often make concessions to this more cynical view. For example, Susan Mendus (1988: 343), who argues that politicians can reconcile the dispositions of the "serpent and dove", says that politicians must sometimes "be prepared to do disagreeable things"; that even though politicians should be "good enough[...]for any world", they will not count as angels.

Common to all these views is the idea that, within the political realm, effective and appropriate political action usually requires accruing some dirt, in terms of both actions committed and characters cultivated. These acts are both all-things-considered (ATC) morally

justified *and* do not result in the violated rules losing their moral force and evaluative strength. While much dirt results from sad contingencies that can be remedied or prevented, much of it represents something more enduring and endemic within politics, including within the liberal democracies that are the focus of this thesis. Even the best form of politics requires some degree of dirt, in the sense that it requires a real and regrettable departure from what morality usually demands.

However, what is *meant* by dirt? And how can something that is permissible, or even morally right, carry with it this kind of polluting moral remainder? In this chapter, I address these questions by elucidating the core idea of moral dirt. I provide a basic definition of moral dirt, one that is compatible with but does not require said dirt being ATC-justified, and one that passes the test of conceptual clarity and cohesion. In doing so, I hope both to shed greater light on some of the commonalities and differences that already exist within the literature, and to create a logical starting point for the more in-depth defences and explorations of dirt that motivate the rest of the thesis. Once this core concept has been defined, we can distinguish instances of dirt based on both severity and on their source/character (as I propose doing via the distinctions between DE, DR, and DC).

First, I will argue for the reality of moral conflicts and that the general structure of moral choice creates conceptual space for dirt. These moral conflicts can entail dirt when a 'correct' option (whether merely permissible or obligatory) is chosen.[7] Recall the five examples from the introduction. These examples all contain a kind of dirt as I define it: even when ATC-justified, a real moral residue results and the agent is, in a meaningful sense, morally different

---

[7] In short, I agree with Goodwin (2009) that it is important between wrongdoings that are merely wrong; wrongdoings that are permissible (or justified) yet not obligatory; and those wrongful acts that are simply wrong. For the sake of the theoretical framework I am drawing within this thesis, however, whether an act is obligatory or merely permissible does not affect the overarching way in which I discuss ATC-justified dirt. When it is relevant, I will discuss it in more detail. For example, when I introduce a tripartite model of DH, this distinction becomes more important in that DE acts tend to have some compelling moral reasons whereas DC and DR acts are often in pursuit of permissible political projects.

because of their choice. Negative moral emotions about such feelings are not pathological but reflect something important. I then distinguish ATC-justified moral dirt from other moral phenomena and explore what it means to be morally dirtied.

Finally, I will examine how inevitable dirt is within politics: firstly, by looking at the question of symmetry, and whether an agent faced with the prospect of accruing ATC-justified dirt can 'keep their hands clean' by avoiding moral pollution altogether. Secondly, I will examine how ATC-justified dirt can arise because of moral luck, thus showing how dirt cannot be easily avoided yet simultaneously remains a valid subject of moral evaluation. In Chapter 2, I will defend this conceptualisation of moral dirt further, including by describing its utility.

1. The Reality of 'Dirtying' Moral Conflicts

The "provocative and [seemingly] paradoxical" (Walzer, 2004: 33) claim at the heart of these discussions is a controversial one: that an act can be both wrong and right. The claim is that politicians can be morally justified in committing real moral wrongdoings. This idea has been dismissed as a rank contradiction at odds with the fundamental premises and logic of moral philosophy (for example, see Nielsen, 2007): how can an act be the right thing to do yet still morally tainting? How can the breach of a moral rule warrant disapproval when said breaching is justified? Depending on how one views the ontology of moral 'oughts', this contradiction seems insoluble and means that, at a metaethical level at least, ATC-justified dirt is a nonsense. It might still be a useful concept practically speaking, but its role is that of a useful fiction. While I certainly do believe the concept of moral dirt *is* useful, it is not a fiction, nor is it incoherent. In this section I will defend the claim that moral conflicts are real and that in some of these conflicts, even the correct option can be seriously dirtying. I endorse the value-conflict

approach to moral conflicts as the most plausible way to defend this claim, building on the account of Michael Stocker.

In short, I view moral conflicts as arising when there is a clash between plural and conflicting values. This account is proposed by Michael Stocker (1990), whose position is defended and expanded upon by De Wijze (2005; 2007), Goodwin (2009), and Nick (2019a; 2019b). Williams (1973a) makes a similar argument in many respects but does not directly link it to DH or distinguish between dirtying and non-dirtying cases. Value conflicts emanate from "conflicting demands on our practical moral reasoning" (De Wijze, 2005: 457), reaching the level of dirt when we "commit a moral wrong in order to do what is morally required" (De Wijze, 2013: 879). "When plural values clash, whatever option we choose we will inevitably have to *violate* one of the values" (Nick, 2019b: 86, emphasis my own), where even though we have decided that one value takes precedence in the context we find ourselves in, the unchosen values retain their "moral relevance". This value-conflict view is ultimately rooted in the belief that moral theory must tell us more than just "what is to be done and how to avoid culpability" (Stocker, 1990: 34), and instead robustly facilitate moral description and evaluation. Under this view, hard moral choices are more than puzzles to be solved, "where all we have to do is find the right answer" (Goodwin, 2009: 19). Reducing them to puzzles of action-guidance, as many have done, is to neglect these more discursive functions.

1.1 Double-Counted Impossible Oughts

Key to the value-conflict view is the idea that these values can retain their evaluative importance even when deprived of their action-guiding function. The importance of not violating a key moral value is retained even if violating the value is ATC-justified — this is how a moral remainder is generated (and how value-conflict views can explain moral dirt).

Moral conflicts arise when we encounter "double-counted impossible oughts". Stocker rejects the notion that ought necessarily implies can. A strict application of that principle, he argues, leaves us unable to properly describe situations in which we are, whether by physical or moral necessity, forced to commit a wrong. By erasing the ought from these scenarios, Stocker (1990: 12) contends, we will find ourselves struggling to describe the harm inflicted or the inherent imperfections in the outcome.

Similarly, Williams (1973a: 181-183) had earlier criticised the view that moral oughts *never* conflict with one another by describing cases of double oughts, where fulfilling both oughts is contingently impossible. When separate obligations exist to do both *a* and *b*, these are separate and independent oughts. We cannot "get off the moral hook" to do *a* by doing *b*, even if *b* is the more pressing ought and renders *a* impossible. Even *if* ought implies can, then, moral conflicts remain. There is no single ought to do *a and b*, as such an ought would be impossible. However, the separate oughts to do *a* and to do *b* can both be satisfied on their own terms. These oughts are both, on their own, possible, yet taken together conflict. For Williams (1973a: 175), this inflexible insistence on 'ought implies can' is based on an attempt to apply the structure of factual beliefs to moral ones. When two factual beliefs conflict, then at least one should be changed or rejected. This specificationist strategy, however, is unsuitable for moral claims (Williams, 1973a: 177-179).

We thus arrive at the notion of impossible oughts. Let us now turn to double-counting. Acts have costs. With some justified acts, we can "double-count these costs: once in taking them as costs and once in holding that despite their being costs, what they are costs for is justified" (Stocker, 1990: 15). If we consider costs solely as a cost, then they are never worth enduring, but if we ask what they are considered a cost *for*, then it can be "better had than not had". This goes beyond mere cost-benefit analysis by accepting that this additional disvalue lingers. Paying a monetary cost for a meal, for example, is a case where we just count the cost

'once', in terms of deciding whether the meal is worth it or not (Stocker, 1990: 16). Sometimes we count the costs twice, however: a ship's captain jettisoning cargo to save the crew's lives is doubtless making the right decision (and far from a dirtying one), but the cost of the jettisoned cargo lingers. Similarly, Williams (1973a: 171-176) argues that when two moral beliefs or two moral obligations collide, the two conflicting beliefs can both retain their value independently and the resulting loss can be rationally bemoaned. Sometimes the lesson from a moral conflict is not that we were inadequately prepared for the choice, but that "both oughts *do* apply" (Williams, 1973a: 177). Following one ought does not require the wholesale rejection of the other (Williams, 1973a: 185). Thus "moral conflicts are neither [all] systematically avoidable, nor all soluble without remainder" (Williams, 1973a: 179).

Double-counted impossible oughts create the need for "non-action-guiding act evaluations", where action-guidance and assessments of wrongness surrounding a particular act come apart (Stocker, 1990: 15-17). The values at play are still of course action-guiding in the sense that, even if they were not decisive, they had an impact on how we decided to act, yet Stocker argues that they retain an evaluative purpose *after* and *outside of* their impact on guiding actions. Evaluative functions can include identifying moral dirt (De Wijze, 2005: 456). If forced to choose between *a* and *b*, when we arrive at an ATC-judgement of what is right, the losing ought is rejected only in a "fairly weak" sense (Williams, 1973a: 183-184). We can distinguish between the force that these oughts have at a contributory level, as inputs into a decision, versus the force that these reasons have at an overall level (Dancy, 2004: 15-16). Just because a reason has defeasible power in terms of its contribution to ATC-action guidance, that does not *automatically* mean it loses all power in terms of evaluation. An ought can be justifiably breached in a way that does not cancel out or amend said ought, along the lines of justificatory defences in criminal law: the defendant concedes that an offence was committed,

yet it was justified under the circumstances.[8] The person who commits violence in self-defence or steals a loaf of bread to avoid starvation need not deny that the offence actually took place: instead, they argue that it was justified.

Allowing for non-action-guiding act evaluations strengthens the *comprehensiveness* of deontic constraints and avoids an immediate fallback to specificationism in the face of conflict. As Goodwin (2009: 164) explains, the kind of 'ought' relevant in DH cases, rather than representing a "conclusion of a moral deliberation" about what to do, instead embodies the inherent categorical wrongfulness that comes from violating *pro tanto* duties. These violations, Goodwin observes, usually entail "[acting] contrary to the intentional harm and separateness of persons arguments." Understood in this way, an act being both rightful and wrongful is coherent. It is these non-acting-guiding evaluations which allow value-conflict theorists to address what Walzer (2004: 33) calls the "provocative and paradoxical" claim that an act can be both right and wrong (or, as Goodwin [2009: 16-17; 144-147] puts it, as containing elements of both rightful and wrongful behaviour.) A moral obligation is breached, with all the significance that this entails, even if the responsible agent can be excused or described as faultless.[9]

1.2 When an Impossible Ought Leads to Dirt

Non-action-guiding act evaluations, in addition to capturing cases of moral dirt, also capture many other non-dirtying value conflicts such as opportunity costs. Sometimes we might bemoan having to give up one valued good or goal for another, such as when choosing between

---

[8] I am indebted to Kenta Sekine for many of the points made in the latter half of this paragraph.
[9] Again, I must thank Kenta Sekine for this point. His defence of faultless moral conflict addresses similar issues to my own.

an educational opportunity or an exciting new career prospect. There is something of value lost regardless of what option we pick, yet this is not dirtying. This is not because these cases lack moral importance. To call such decisions dirtying is not only terminologically excessive, but also deprives dirt of its distinctive conceptual potential, leaving DH a mere synonym of value conflicts more broadly. An act, whether straightforwardly wrong or ATC-justified, needs something else to be dirtying rather than solely entailing the breach of a non-defeasible ought.

DH cases arise when we are dealing with values that are important enough that they cannot be traded off against one another "without a significant remainder" (Nick, 2019a: 2). Stocker attempted to specify this further by saying that dirtying acts are those that involve "people being wronged, they and their trust, integrity, and status as ends are violated, dishonoured, and betrayed; innocents are killed, tortured, lied to, deceived" (Stocker, 1990: 17).[10] It goes beyond merely giving something of value up, but must meet one of these disjunctive separate criteria. Acts that are morally dirtying exist outside of DH conflicts (Stocker, 1990: 26). What is different about the dirty acts in these scenarios is only that they are, in these cases, ATC-justified. In DH scenarios, these acts are "1) justified, even obligatory, but (2) none the less somehow wrong" (Stocker, 1990: 10). They retain their "disvalue" even when done as part of "a justified, perhaps obligatory, whole" (Stocker, 1990: 13).

This definition of dirt is suitable for the more traditional DH choices such as the TBS (though in need of further specificity), but is there space within this view for instances of ATC-justified wrongdoings that are not discrete decisions to commit determinate wrongs for determinate benefit? The *basic* structure of choice created by Stocker's account creates space for such an endeavour: dirtying conflicts are those where the required violation creates a moral remainder and dirties the responsible agent's moral character, and where the "non-action-

---

[10] I also doubt that all cases of "people being wronged" and innocents being "lied to" or "deceived" can be placed within either an episodic framework or inflict the same severity of dirt.

guiding act evaluation" is of a particularly harsh or serious kind. An act that is "morally unavoidable" in the circumstances in which an agent finds themselves is still "said to tell against act and agent" (Stocker, 1990: 12). The notion of a justified yet simultaneously "somehow wrong" act seems to capture what is at stake in cases of permissible political dirt. Stocker's attempt to further specify what makes something dirtying is less helpful here, since we can think of examples of other agents who are plausibly dirtied that his criteria does not cover: agents who are corrupted, who misbehave in less dramatic ways, who become part of and identify with morally compromising structures. These are cases where that disvalue falls short of the threshold Stocker sets out. A broader definition of dirt is needed.

These cases cannot all be viewed as mere opportunity costs. Sometimes there is a cost intrinsic in the act itself; a strong moral ought is broken even when part of a justified whole. A choice between valued life-plans, despite its importance, need not generate dirt, but if one of those choices is to require movement away from a 'moral baseline' — some kind of moral violation or degradation that is lamentable on its own terms — then dirt does arise. Suppose one of the valued life-plans requires giving up on one's established relationships, or abandoning a virtue that has previously defined one's entire life. For example, imagine a soldier abandoning the front line of a just war to care for his family or vice versa. Such dirt is not akin, in degree or in structure, to the dirt created by scenarios such as TBS, but moral residue still falls upon the agent. To call that residue 'dirt', and then to distinguish *between* those different types of dirt, seems to me the most effective step. I will expand upon these different kinds of dirt later in the thesis. For now, I hope to have argued that ATC-justified dirt is a coherent concept.

2. What is Moral Dirt?

What is this moral dirt that we can, sometimes, justifiably accrue? Our examples are diverse, and some risk being shut out by standards such as Stocker's, despite this shared foundation of ATC-justified wrongdoing. This section will outline a more 'basic' definition of dirt, one that captures any kind of noticeable moral taint that might be justifiably incurred, from the mundane to the extreme. From there, we can delineate between different categories of political dirt in greater detail (later breaking this category down into grimy, dirty, and bloody forms of moral transgression, in addition to categorising different ways in which the choice can present itself).

This task requires setting out the necessary and sufficient conditions of dirt. In and of itself, to say something is dirtying does not say all that much about either the act or its broader circumstances. It does not tell us whether the act was justified or excusable, what kind of impact it has on the responsible agent's character and that of their broader polity, what sort of reckoning or accountability is called for, or whether such dirt was inevitable or inescapable, amongst many other questions. The focus in this thesis, of course, is on those acts that are ATC-justified, but this is an extra criterion, beyond the core concept of dirt.[11] As necessary (and jointly sufficient) conditions of dirt, I propose:

1. A moral 'ought' is not acted upon, grounding an evaluation of the act that *need not* be action-guiding.

---

[11] Definitions of moral dirt in the DH debate have focused on what dirt means in the very specific context where it is ATC-justified. I move beyond this focus because I believe that there are multiple conditions that can be said to render an instance of dirt ATC-justified. There is no single definition of ATC-justified dirt, then, even if we can, as I think we can, define moral dirt more generally. This discussion helps to address what can be cited as an ambiguity in the DH framework: Coady (2018: 2) argues that sometimes the term DH is used "for any conspicuous immorality as well as the more interesting and technical sense we are concerned with". Far from this being a bug, I believe it is a feature. It thus makes perfect sense when Walzer (1973: 174) speaks of how it is "often right" to incur dirt, and Levy (2007: 52) notes how dirty actions are "part and parcel" of everyday politics. It can be tempting, easy, or advantageous to perform a dirtying act, and so too can it seem that it is simply a matter of political reality. This does not mean that it is justified to do so. An act can be dirtying without creating a DH conflict that is of moral interest. Drawing out the distinctions between different categories of dirt as I do in this thesis helps to evaluate the extent to which a specific instance of dirt is justified.

2. There is something categorically wrongful about the act itself and/or the character of the responsible agent. Something morally deficient is revealed and legitimate moral criticism can be made as a result.

3. There is a moral remainder, reflecting an uncancelled disagreeableness.

These conditions are shared by DE, DC, and DR, which are distinguished from one another by their own additional necessary conditions. The respective scopes of these different types of moral conflicts are delineated by additional criteria that are applied on top of these three conditions, and which will be set out in Chapters 4 (for DE), 5 (for DC), and 6 (for DR).

2.1 An 'Ought' Not Acted Upon

The first condition follows from the account of moral conflict sketched out earlier in this chapter. In short, dirt arises when a moral imperative is not acted upon. This ought is double-countable in that it lingers as an evaluative metric independently of its role in determining what should be done. This condition is necessary, yet not sufficient, for generating moral dirt. I present these oughts as being of a morally important character: not necessarily moral themselves, but if they are non-moral oughts, such as political efficacy, it should still be possible to evaluate their moral importance.[12]

What is the substance of an 'ought' not acted upon? I believe that these are best viewed as resulting from conflicts between competing values, which we can call the value-conflict (VC) framework. The primary alternative to this view in DH debates is to view these conflicts as instead arising between competing moral frameworks when incommensurable kinds of

---

[12] Thus I do not completely reject the view that DH cases arise "when morality clashes with some other rational necessity of a profound kind that correctly overrules it" (Coady, 2018: 3). The prospect of a moral norm clashing with another rational necessity, and thus generating permissible dirt, is allowed under my account (yet does not result in the complete suspension of morality). However, I do not treat a clash between moral and non-moral oughts as a necessary condition of DH because I want to allow for conflicts *within* morality.

imperative clash. The right thing to do according to one of these frameworks might well be the wrong thing to do according to another. There are two prominent 'competing-framework' views: consequentialism-versus-deontology and public-versus-private.

The first is to argue that the competing demands of consequentialism and deontology (hereafter the consequentialism-versus-deontology, or CvD, view) generate uncancelled and potentially dirtying demands. This is how Walzer (1973: 161) presented DH: an act "may be exactly the right thing to do in utilitarian terms", yet still entail committing a moral wrong and running afoul of "absolutist" moral prohibitions. Walzer regards both deontological (usually absolutist) and consequentialist (usually utilitarian) schools of thought as "important but potentially irreconcilable strands in our moral thinking" (Coady, 2008: 79). The Walzerian politician, then, is "torn between [her] deontological and consequentialist commitments" (Nick, 2019a: 2-3).[13]

The second approach is to say that these conflicts result from a clash between the requirements of public and private morality (hereafter PvP). Dirt arises from doing something that is "right" by the standards of public morality (such as fulfilling the obligations of one's role), yet "wrong" by the standards of private morality (such as integrity). Different roles purportedly demand allegiance to different moral theories, hence the conflict. This view assumes that DH are exclusive to politics, or at least to politics and only a few similar public domains. Thompson (1987: 11), for instance, characterises these problems as arising for "the political leader who for the sake of public purposes violates moral principles", where dirtying actions can be justified from the point of view of the polity but not from the individual agents authorising or implementing them. Nagel (1993: 79-84) argues that, though public and private morality flow from the same source, they each manifest in differing ways and therefore create

---

[13] The CvD approach is far from exclusively Walzerian and is often treated as *the* DH approach by both its critics and supporters (for example, Klockars, 1980: 35-36; Ramsay, 2000: 30; Nielsen, 2007: 30; Meisels, 2008: 152-154; Bellamy, 2007: 510; Thalos, 2018: 171; Smith, 2018: 205).

different prescriptions. Individuals exist both as private citizens and as functionaries of the social role that they hold, and both public and private morality exercise a claim on the individual.

The alternative to (both) competing-framework approaches is, as discussed, the idea that dirt can arise when "plural values clash" (Nick, 2019b: 86), creating "conflicting demands on our practical moral reasoning" (De Wijze, 2005: 457). In contrast to the competing-frameworks models, the VC account is more flexible and context-sensitive about what the prongs of the dilemma are and thus more ecumenical about the scope of choices that can be analysed within it. The values in question *can*, but need not, take the form of deontic constraints and duties, consequentialist considerations, or attachments to various roles and public duties. Furthermore, such an account does not *insist* upon the creation of a bespoke political role morality that "somehow transcends the general morality" (Coady, 2018: 4). Granted, a VC account is not incompatible with such a political role morality, but nor does it require committing to this "political exceptionism" that Coady argues is "inherent" in DH theorising (ibid).

The chief objection to the CvD and PvP accounts is that they fail to capture many cases that share pertinent and relevant attributes with paradigmatic examples of DH. The additional necessary condition of competing frameworks "unduly limits the kind of situations" we can analyse (Nick, 2019a: 3), with no countervailing benefits to justify these exclusions. CvD accounts cannot, by their nature, capture conflicts between deontological constraints or values. Conflicts *within* deontology are precluded. Clashes between different loyalties, rights, and ideals do not necessarily rise to the level of dirt, and often deontic constraints are defeasible or *prima facie*, but on the occasions where deontological norms *do* generate dirtying conflicts, CvD accounts are silent (or must deny that these really are dirtying conflicts). De Wijze (2007: 11) points to the example of Antigone from Greek mythology as an example of this. What

animated Antigone's inner turmoil was how two deontological obligations — her duties to her city and to her family — came apart. That consequentialist reasoning sided with one of them was more of a coincidence than an intrinsic feature of her dilemma. Removing that consequentialist sway does not dissolve the dilemma for Antigone, and both the value that she abides by and the one that she violates are important independently of consequentialism. The "lesser evil" she aimed for was one that had to be justified in deontological terms. The moral remainder that was created, whatever option she picked, was expressed predominantly through deontological language.

Salvaging the CvD account, then, requires either denying the possibility of conflicts between deontological values or manipulating them to conform to the CvD model. One way of doing this might be to argue that deontological values within politics are interconnected, mutually support each other and, drawing from Dworkin (2001: 90), should be respecified to "hang together in the right way". This is an unconvincing specificationist defence. Firstly, it would be odd, and counter to Dworkin's views, to create a space for consequentialist moral pulls, and *only* consequentialism, within an otherwise harmonious deontological moral framework. Secondly, real-world clashes between values still arise even if we accept that they can, in theory, be made to co-exist harmoniously in the future. Thirdly, we might be asked to choose between instances of the same value (Stocker, 1990: 242-249), such as meeting the demands of a given value for one person or another (Hurka, 1996). Attempting to arbitrate between our choices in these instances without introducing new values, and thereby moving from monism to pluralism, is not easy (Stocker, 1990: 246-249). Ultimately, we can reject the CvD account for reasons of scope while remaining largely agnostic on the deeper questions of value pluralism. The existence of *contingent* clashes is sufficient to show that the CvD framework cannot accommodate what appear to be standard cases of dirt.

Similarly, PvP also unnecessarily and unhelpfully limits the scope of cases we can consider. Even if we assume that a stable distinction between public and private moralities can be drawn—itself a debatable assumption—cases of ATC-justified dirt cannot be defined exclusively as PvP instances. Private life could have moral dilemmas where the language of moral dirt can help describe tensions between obligations such as friendship and honesty (Nick, 2019a: 2). Likewise, there are many examples where dirtying tensions exist entirely within public life, such as cases where we must use undemocratic tools to achieve democratic ends (Nick, 2019b: 82-87). Waldron (2018: 229-230) describes cases where a single agent faces conflicting moral obligations that result from different public roles or different political and legal principles, such as when a politician is bound by collective cabinet responsibility to implement a decision they regard as wrongful. The necessary conditions for dirt "lie elsewhere" (De Wijze and Goodwin, 2009: 531), and what generates the dirt is often independent of whether we are acting in a public capacity (De Wijze, 2007: 10). Take the Jim example Williams (1973b: 98-99) uses, where a man can kill one person to satisfy the capricious whims of a brutal commander or allow that person and his nineteen friends to all be slaughtered. Jim's status as a private citizen, along with whether the deaths he can prevent are characterised as public or private consequences, seems beside the point when deciding whether his case counts as dirtying.

Ultimately, for these reasons of scope and accurately locating the necessary conditions, I think the VC view best describes what it means when oughts are not acted upon. It is also worth noting that the VC account is more ecumenical. The VC account is inclusive not just of the cases used by the alternative conceptualisations, but indeed of the conceptualisations themselves. The VC account "essentially sidesteps these different understandings [yet remains] compatible with any [rival] definitions of [dirt] that one finds most congenial" (De Wijze, 2013: 881). To say that ATC-justified dirt should be viewed as arising from a clash between values

does not commit us to saying that this clash can never take place along CvD or PvP lines, but only that they are not *necessary* conditions for DH.

The explanatory power of the VC account can be illustrated in relation to our five examples. Firstly, consider the examples of Edna and Paul. These examples involve morally important elections where winning seems to require some kind of morally unideal action. On the surface, these can fit within the CvD framework. The consequentialist gains of winning the election (or conversely, the losses of losing it) seem to outweigh deontological rules. However, this does not quite capture what troubles Paul and Edna. What troubles Paul, for instance, is not that he is breaking behavioural rules, but that he is cultivating undesirable traits and dispositions. What troubles Edna is not that she herself is directly responsible for the institutional wrongdoings, but that she is exploiting, perpetuating, and participating in them.

Ned and North Takoma's political leaders can face plausible deontological prohibitions against corrupt deals and gerrymandering. However, political concerns need not be exclusively consequentialist: effective political participation and obligations to one's political allies can be expressed politically. More fundamentally, these two cases expose the limitations of PvP. PvP may somewhat capture the loss of personal integrity and the cultivation of unsavoury frameworks in Edna's and Paul's dilemmas (though even these are ultimately of public significance), but with corruption and gerrymandering, both 'oughts' are public in nature. The TBS example perhaps aligns most neatly with the competing-frameworks approaches, especially CvD. However, even if we take TBS as a CvD or PvP conflict, this should be seen as an example illustrating the kind of values that might come into conflict, rather than reflecting the necessary conditions of these conflicts.

For an act to be 'dirtying', additional criteria must be met.[14] The first of these is that there must be something about the act in question that can be used to speak against both the act and the agent responsible for it. There is a moral deficiency—not just the absence of some good quality, but something actively and affirmatively negative. How exactly an act does this will depend on the act in question and the specific values affected. The decisions in the five examples all have something about them that grounds negative commentary against both the act and agent. This can entail participating in an act that plausibly involves a *pro tanto* wrong (torture, gerrymandering, bribery) or showing an inconsistency with the agent's professed moral and political values.

There must be something to the act that "is said to tell against act and agent" (Stocker, 1990: 12). This could be if the broken categorical norms have a "categorical and mandatory force" (Goodwin, 2009: 38, following Gardner and Macklem, 2004: 465). In extreme cases, it can be an act so contemptible, so outside what is normally justified, that for it to be considered "shows a corrupt mind" (Anscombe, 1958: 17). In other cases, the agent becomes shaped by or complicit with the corrupt institutional arrangements around her or the corrupt actions of other agents. The agent is then "voluntarily" co-opted into "committing immoral acts or participating in immoral projects" (De Wijze, 2009: 309), becoming part of a "complex of immorality". The agent thereby forfeits "innocence-as-a-disposition" (Tillyris, 2016a: 169).

Another way this criterion could manifest is if the agent acts in a manner contrary to their political commitments in a substantive and morally important way: Edna does this. Some political values are more than just policy programs. These values place constraints on

---

[14] Here I use the word 'act' in a broader sense that encompasses strategies, practices, and decisions about character cultivation.

individual behaviour too. To violate these is to act in a morally deficient way, contributing to the political problems that the agent opposes, even if the act is ATC-justified. To use Wertheimer's (1987) terminology, this goes beyond merely violating a "phenomenological" baseline. Those that the agent is acting against have morally important rights—rights that the agent accepts—not to be treated in a way that violates their rights, and yet these rights are being violated. Such an act is dirtying for the culpable agent. It demonstrates inconsistency with one's ultimate values, or at least an admission that they cannot be fully actualised in the world as it is.

2.3 Moral Remainders

The final necessary criterion for generating dirt is that there must be some tangible moral residue, something tainting and polluting that attaches to the agent because of the decision. The moral disagreeableness involved in the act is not dissolved and cannot wholly be passed onto others. This dirt tracks some kind of responsibility for a wrongful act (De Wijze, 2018: 130) or some evidence of moral corruption on the agent's behalf. When gaining this dirt is justified, "praise" and "opprobrium" can both be simultaneously appropriate (De Wijze, 2018: 141). There is a "residual moral claim against" the agent (Hall, 2022: 218).

The "remainders thesis"[15] requires more than just pointing to the fact that a moral imperative remains uncancelled. It must represent an uncancelled disagreeableness (Williams, 1978: 62-63), more serious than a forgone opportunity cost. In addition to there being something to be mourned about the path not chosen, there must be something lamentable about the path that *has* been chosen. There are legitimate grounds for others to complain about how

---

[15] As described by Goodwin (2009: 47).

they have been treated or what the responsible agent has become, even if they ultimately accept that the action was ATC-justified. These acts do not lose their morally polluting character and evaluative force even if ATC-justified (Williams, 1978: 71). What is crucial is that this dirt is real. It is more than a mere phenomenological observation or a useful fiction. This is what differentiates the idea of justifiable moral dirt from the evaluations put forward by, say, consequentialists and threshold deontologists.

3. What *Isn't* Moral Dirt?

If a case does not meet these three necessary and jointly sufficient conditions, then it is not dirtying. What sort of cases does this definition exclude? As discussed, not every moral choice that involves a morally significant opportunity cost counts as dirtying. To be dirtying, there must be intrinsic costs to the path that is chosen, dirt that persists regardless of whether or not it is part of an ATC-justified whole. This creates a lower boundary of scope, beyond which choices are not dirtying. Illustrating this lower scope is important to defeat the "enthusiasm to detect the sway of dirty hands" (Coady, 2018: 2) found in parts of the literature. Such an enthusiasm leads to the terminology being applied in a way that both mischaracterises other issues and deprives dirt of its distinctiveness. I hope to avoid this challenge by sticking rigidly to the three conditions. This section will look at some of the cases that are excluded by the definition presented in the last section.

Firstly, there are cases where agents are faced with painful and perhaps even "horrendous" dilemmas, where obligations and duties clash, yet nothing can rightly be said against either the agent, their decision-making, or their integrity and internal virtues. Moral remainders arise, double-counted costs are incurred, and oughts are not acted upon, but the agent incurs no dirt at the point of decision-making. This category includes cases known as

"lifeboat ethics" (Hardin, 1974), wherein limited resources or the natural facts of the world prevent us honouring all our duties or attending to what is owed to each individual. A lifeboat captain deciding who to save, a ship captain jettisoning cargo to keep the ship afloat, and choices about distributing insufficient medical supplies are all such examples (De Wijze, 2007: 13-14). Lifeboat ethics cases are difficult but not for the same reasons as ATC-justified dirt.

Provided there is no wrongdoing in the selection process and the agent bears no responsibility for the shortage, no dirt is inflicted (and such dirt would be hard to justify besides). The very act of choosing in these predicaments is not dirtying. There might well be additional factors that make the decision dirtying, but the lifeboat ethics framing itself does not entail dirt. It might be that entering a career where these decisions are to be expected or mandatory might require adopting certain dirtying dispositions or engaging with certain dirtying institutions. However, the decision *itself* is not dirtying, even if it is harrowing and even if there are serious uncancelled costs. Any such dirt here is instead more of the kind associated with DC and DR.

Some value conflicts *are* the responsibility of the agent, and *do* say something about their moral character, yet do not reach the level of dirt. De Wijze (2007: 13-14) gives the example of university administrators who, fearing that an invited speaker's attendance would cause disruption, decide to block the speaker, opting for security over freedom of speech. De Wijze describes this as a value conflict that is not dirtying due to it not involving duplicity, deceit, injustice, or an absolute denial of these core values. Granted, the example as given is underspecified: what affirmative right did the speaker have to the slot in the first place, for instance? What is the broader impact of the decision on freedom of speech within the academy? It might be that, under certain circumstances, this move is dirtying. However, at its core, it does not meet either of the three criteria: it does not necessarily entail the violation of whatever 'oughts' the administrator has to promote and protect the relevant values (as it is unlikely the

administrator has the sole responsibility to maximise both values at play). As a result of this, it does not necessarily indicate something morally reproachable about the deciding agent and does not necessarily carry with it a substantial moral remainder.

A more useful reconstruction might be broader policy choices. Plausibly a university administration must balance different considerations in its policy for accommodating invited guest speakers. Some of these values are, arguably, not *directly* related to morality, such as cost, popular demand, and capacity. Others, such as security and freedom of speech, are morally important. Trading these values off against each other will certainly have moral costs as the administration will be falling short in living up to them. However, they *may* not be violating any of them to such a degree as to manifest dirt, particularly if they all remain above whatever threshold is deemed necessary. An important value conflict is not sufficient to ground dirt.

Sometimes the act in question is presented as a categorical *pro tanto* wrong, and therefore it is dirtying even when ATC-justified. The designation of *pro tanto* wrongness can be debated. Sometimes this takes the form of acts being falsely described as dirtying. Take lying or, more specifically, embellishment in storytelling and fiction; or lying to stop an axe murderer or Nazi from finding their victim (Coady, 2008: 117). These acts are dirtying *iff* one believes the prohibition against lying is absolute (including if only for evaluative purposes). However, I think it is perfectly plausible to argue that the immorality of lying *is* cancelled in these cases and that no compelling account of the virtue of honesty can condemn them. There is no 'ought' that is ignored, no moral remainder arises, and nothing compelling can be said against what the agent is or has done. This is true with most value choices where the only thing of value given up is the opportunity cost of the path not chosen.

Sometimes acts are only *prima facie* wrong. DH, in contrast, covers those types of wrongdoing that are either not cancelled by the circumstances or are non-defeasible altogether

(*pro tanto* wrong). So, while it is usually wrong to inflict pain without consent, if we must perform life-saving emergency surgery without anaesthetic, that wrongness is set aside altogether, rather than merely reduced to a non-action-guiding evaluation (Levy, 2007: 39). A similar case might be found in the ethics of self-defence: it is wrong to harm someone, but this is a *prima facie* constraint, and proportional harm in self-defence is not seen as dirtying in most cases. Identifying whether an act is dirtying or not requires first asking whether there is wrongfulness attached to the act, or whether there is merely a *prima facie* prohibition, or whether there are other reasons why the dirt does not stick. In short, once "properly understood and contextualized", certain rules might not be dirtying once broken (McDonald, 2000: 188-189).

In other cases, the language of dirt might be applied erroneously to categories that are morally disturbing for their own reasons (Coady, 2018: 2). In addition to cases of "lifeboat ethics", there are cases of "spattered hands" (Rubenstein, 2015). Suppose a humanitarian organisation finds their aid is being seized and then abused by malevolent actors, such as gangs taking food supplies and using them as leverage in warfare against rival factions. Aid agencies are then left with the dilemma of whether to withdraw aid. Rubenstein argues that it is erroneous to use the language of dirt to describe these predicaments. Instead, "spattered hands" dilemmas are the proper term for these cases. No wrongful action is being taken by the agency in question and yet their good deeds are being perverted, predictably and yet in a difficult to prevent way, by other actors. This is assisting another agent's evil project, but crucially the aid agency is not a willing participant in the wrongdoing and is not consciously trying to bring about a "lesser evil". The phenomenon, then, is something different. Spattered hands cases can, I contend, become dirtying (most likely in the sense of DC or DR) if the agency begins incorporating the furthering of the gang's goals into its evaluations or starts using this corrupt economy more proactively to deliver aid. Ultimately good deeds being perverted, even

predictably, into negative consequences by other actors is not sufficient to generate dirt for the agent performing the good deeds. Something more needs to be done to demonstrate that a moral ought has gone unmet or that whatever dirt arises from this scenario falls on the responsible actor.

Another example is Sophie's Choice, the notorious dilemma that the character Sophie faces in the novel and film of the same name. Sophie is forced by a Nazi guard to choose between her two children. One will die, the other will be saved from certain death. If she does not choose a child to let die, then both will be killed. De Wijze (2007: 13-14) lists this as an example of DH. By extension, whatever choice Sophie makes is dirtying. I think this is mistaken. To be sure, Sophie faces a dilemma (in the form of which child to save). De Wijze (2007: 9) uses Sophie's Choice to defend the idea of dilemmas, as opposed to exclusively conflicts where there is one right answer, also generating ATC-justified dirt. I agree with him on this conceptual point. However, I do not think that Sophie's case satisfies the second criterion for an act to be dirtying. This horrific situation says nothing against her character. The appropriate language for what has befallen Sophie, then, will not be moral dirt, but instead moral injury (Kellison, 2021: 447-450), violently forced upon her, deliberately inflicted as a core part of the villain's goal to degrade and dehumanise her. Cases of moral injury, which are potent phenomenological experiences yet need not track moral reality, should be strongly distinguished from cases of moral dirt, even though they are related and potentially overlapping concepts.

4. Can Dirt Be Avoided?

Throughout this chapter I have advanced the claim that ATC-justifiable moral dirtying is a coherent moral phenomenon. Effective political action will, for most actors, require some

engagement with morally dirtying practices, dispositions, and crafts. Dirt thus becomes inevitable. Adding to this, political actors can sometimes be confronted with further scenarios where even the morally abhorrent becomes ATC-justified. While I do not want to argue that dirt in politics is always inexorable and that politics entails leaving morality at the door altogether, the goal of making this more nuanced account is best served by breaking down the concept of ATC-justified dirt rather than denying it completely.

Avoiding dirt is not a simple task. I demonstrate this firstly by examining how certain dilemmatic choices are symmetrical, in the sense that dirt cannot be avoided once confronted by them. Secondly, I examine the role moral luck plays in determining where dirt falls and in what quantities, and to what degree it is ultimately justified. Finally, I discuss the concept of "many hands". This is the idea that responsibility for a decision can be diffuse in complex political systems, but that personal responsibility remains important. This helps illustrate another way in which political actors cannot avoid dirt: they are contributing to a great many decisions.

Before that, though, a couple of preliminary points. Firstly, I reject the notion that we can simply say, as a rule of politics, that we can *always pretend* to avoid dirt. In the next chapter I will discuss in more detail how minimising exposure to dirt is useful and, *ceteris paribus*, desirable both for positive reasons (maintaining a claim to exemplarity, or at least some lesser degree of admirable virtue) and negative reasons (avoiding the risks of corruption and maintaining the force of deontic constraints). While it is possible, and indeed common, to try and accrue these instrumental benefits by faking these virtues and wearing "clean gloves" (Bellamy, 2010), some of these benefits do come in the form of intuitions, instincts, and dispositions. These cannot be quite so easily faked. Much as the cunning politician spots opportunities and dangers that their more innocent counterparts cannot, 'cleaner' actors can offer unique perspectives and a different kind of moral imagination. They are also less

susceptible to the risks of corruption, as they are strongly disposed to avoiding unnecessary dirtying acts, rather than these acts (merely) being the result of a more calculated decision.

Secondly, as demonstrated by the distinction between different forms of dirt that will be discussed in chapters 3 through 6, even if we adopt a certain view of the relationship between morality and politics in one area, it does not mean that we are bound to do so in others. Take the (admittedly contested) idea that compromise is essentially dirtying, and that, as an essential component of politics, therefore politics is dirtying too. Let us assume for the sake of argument that this is true. If so, it follows that there is *a* departure between morality and politics. What does not follow is that this departure itself is without moral implications, or (crucially for the point here) that this completely sums up the relationship between morality and politics. It is tempting to extrapolate from, say, the TBS or any other example and generalise about how politics and morality relate to each other (or not). This would be misguided.

4.1 Symmetry

The presence or absence of symmetry is a characteristic of conceptualisations of DH first made explicit by Nick (2019a): can a politician emerge from a DH conflict with clean hands? Applied more generally to dirt, the symmetry division asks whether an agent can avoid dirt altogether when faced with a scenario where the ATC-justified *right* action entails dirt?[16] This question is important in terms of ensuring the proper allocation of moral responsibility, the correct description of an agent's phenomenological experience and moral character, and the right understanding of what is at stake for the actors making the decisions.

---

[16] It is important to note that often this debate has not been conducted explicitly, but instead has been a difference between theorists that, barring a few exemptions such as Nick, has often gone without comment.

Under the asymmetrical view, keeping one's hands clean is possible providing one adheres to deontic constraints, abides by the demands of virtue, and eschews even the justified violation of these important moral norms. If a theorist talks of refusing to dirty one's hands, as Walzer (1973: 165), Waldron (2018: 225), and Dovi (2005: 133-134) do, they are at the very least adopting the language of asymmetry. They acknowledge that there will be real moral costs that emanate from refusing to dirty one's hands, but the agent keeps their hands clean (even if doing so is ATC-regrettable). What does it mean to have clean hands in a scenario such as this? Various answers have been proposed within the literature: maintaining innocence or integrity, avoiding direct causal responsibility for wrongdoings suffered, or abiding by given deontological rules (Nick, 2019a: 7-10).

Many of these accounts assume a certain primacy of deontology over consequentialism, or private over public morality. These accounts assume that deontology/private morality are the routes to moral cleanliness. This is an unsustainable assumption if we instead adopt a VC understanding of dirtying moral choices, where one value cannot cancel out the disvalue caused by violating another and fulfilling one ought does not cancel another. Similarly, it is unclear *why* no dirt can attach to the dictates of consequentialist or public moralities. This prioritisation of one kind of morality over another requires a justification that is rarely offered (Sorell, 2015).

Under the symmetrical view, by contrast, dirt is unavoidable. Once one is faced with a DH conflict, it is too late to emerge with clean hands. It becomes impossible to maintain moral innocence, purity, integrity, blamelessness, or however we might want to describe cleanliness. Hollis (1982: 394) argues that, when faced with a DH dilemma, "it is too late for clean hands, whatever [the DH agent] does", and no action, not even resignation, can allow the individual to stay clean. If dirt can flow from breaking duties and obligations (De Wijze, 2007: 4), or any intentional infliction or allowance of serious wrongdoing (Yeo, 2000: 157), then there is no

dirt-free way out of these predicaments. Whatever we choose, an obligation is broken, or wrongdoing occurs for which we were in some way responsible.

Of course, it is possible that there might be some symmetrical conflicts and some asymmetrical conflicts: ultimately adopting this position would still require pushing against the asymmetrical intuition that there is one side of the dilemma that is uniquely dirtying and one reaction that we can always associate with cleanliness. For a conflict to be symmetrical, then, there is no way for an agent to avoid an outcome that generates *all* three criteria of dirt set out earlier.

It is also worth noting that the term 'symmetry' is potentially misleading, in that (in the technical sense used here) it does not imply that all options are *equally* dirtying. Instead, it is symmetrical in the sense that there is no *clean* option, even if there is a comparatively cleaner one. The moral remainders created are not of the same strength or character, yet both prongs of the dilemma create tangible moral dirt upon the agent. With regards to life choices where all options are permissible, yet some are uniquely dirtying (such as how I describe DC), asymmetry has greater purchase.

4.1.1 Symmetry and Dirty Hands

Symmetry seems especially plausible in cases where the stakes are very high, and the moral pressures are serious. Take the TBS. Dana has an obligation to protect the innocent lives under her rule from terrorist attack. Eschewing the chance to do something that is so clearly a lesser evil (as the traditional formulations of TBS assume, though both assumptions can be questioned) to save those lives can plausibly be described as dirtying. It satisfies all three of the criteria outlined previously. A morally significant responsibility has gone unfulfilled, and obligations to one's citizens have gone unmet. Dana has failed in the duties that she accepted,

and a moral remainder and need for accountability arises. As Finlay (2011) has argued, even if we believe that the ATC-justified option is to not torture, dirt can arise from prioritising the culpable terrorist's right not to be tortured over the right to life of countless innocents.

An asymmetrical view, in contrast, requires "some act of moral compartmentalisation in which [the agent] detaches her conscious decision to choose one course of action over another (e.g. not torturing) from the consequences of that decision (e.g. the death of innocents)" (Nick, 2019a: 11). This compartmentalisation is "both unlikely and undesirable", failing to capture the causal role our decisions have and the unavoidable repercussions on innocence and integrity that come from DH dilemmas. We cannot fully grasp the stakes entailed by DH scenarios unless we acknowledge the moral stakes and the moral duties that go unfulfilled and violated on both sides of the conflict. If one side did not inflict dirt, then we would not be in the realm of choosing lesser evils.

The asymmetrical view might have greater purchase as a phenomenological explanation, particularly in cases where external immorality is the igniting factor. Return to the TBS for a moment. Dana may more pointedly feel the dirt that comes from authorising torture. Dana's causal role in authorising torture is a direct causal role and what it says about her character (that she is not the unequivocal opponent of torture that she imagined herself to be) is clear. The dirt that comes from letting the bombs go off is, in contrast, a more distant dirt. The abject immorality and principal responsibility, after all, lies with the terrorist. Yet Dana can prevent those harms and has a duty to do so.

In Dana's case I ultimately believe that the symmetrical view best captures the phenomenology of DH as well as better following from the VC account's more general structure of choice. It would be implausible for Dana to argue that she left that scenario with clean hands if she had done nothing. She could have prevented the harms and had the ability to do so with, as the TBS example usually stipulates, reasonable degrees of certainty and

proportionality. Actors motivated by moral considerations cannot just ignore their failure to stop these horrors. They will feel guilt, remorse, or some kind of moral responsibility. They will be asked to account for their choices and to remedy the costs of their (in)actions where possible. After all, serious evils and great violations of values happened that they could have prevented. Moral dirt is the appropriate way of describing the moral residue left behind in such cases. This seems like a plausible understanding of both the external and internal assessments of their actions. Actors like Dana are "caught in a cleft stick", according to De Wijze (2009: 316), "morally culpable if they do and morally culpable if they do not."

4.1.2 Symmetry and Other Kinds of Dirt

Is asymmetry true in less stark cases? With regards to choosing whether to follow a path that requires DC or engaging with DR, the choice seems to be more asymmetrical.[17] In some cases, standing by and doing nothing, or acting with additional constraints that hamper one's effectiveness, can generate moral dirt, in that this prioritisation of one's own cleanliness morally speaks against the agent. Symmetry with regards to how deeply an already-committed politician embraces this dirt seems more plausible: if one's political effectiveness is attached to morally important goals or others are counting on one's success, to hold back appears to be a symmetrical choice. This is especially so if one was in a unique or privileged position to act. People will be entitled to ask why *this agent* holds such a position instead of a more effective actor. The reasons for disapprobation are political, but political criticisms can and usually do have some moral weight. However, at a prior level, the choice to enter politics and to adopt the unsavoury dispositions and counterintuitive virtues of politics is a choice that can be freely and

---

[17] With regards to less severe episodic cases, my view is that symmetry holds *if* the agent has a duty to prevent the greater wrongdoing that they are faced with.

permissibly made. We do not judge people for refusing to enter politics, any more than we judge people for refusing to enter other morally valuable careers such as medicine, education, or care.

Now, this argument does presume that no dirt arises from abstention. In certain cases, this might be disputed. If there is gross injustice going on, to be neutral and prefer a solitary and apolitical existence might be wrong. Sometimes it might be wrong to just walk on by and continue one's happy apolitical existence. When I talk about DC, then, I assume that the person making the decision has no special responsibility, such as one generated by confrontation or unique ability, to delve into the political arena. Non-political actors might *still* have a responsibility to avoid perpetuating harm (which might have implications for the dispositions that must be cultivated), but a wholesale commitment to pursue democratic politics as a career is usually a choice.

Sometimes symmetrical language does arise within the context of what I would call DC, where a focus on the background dirt of politics is what is said to make dirt unavoidable. Hampshire (1989: 170) describes politicians as needing to forever "be prepared for the occurrence of an uncontrollable conflict of duties in situations which seem to exclude the possibility of a decent outcome." "The experience of political power" is, for Hampshire (1999: 72), "the experience of unplanned responses to emergencies in constant succession." In handling such emergencies, the political leader must prioritise "their [people's] lives and their freedoms" ahead of "his own integrity and moral respectability." This captures how, once the DC decision has been taken, it is impossible to maintain completely clean hands. The political actor agrees in advance to regard, within eliminative episodic contexts, their own moral integrity as secondary to the basic needs of their citizenry. Describing this decision as dirtying seems to best recognise what is at stake.

What this initial decision does *not* do is account for all that episodic dirt that might follow. DC does not change the moral character of an agent so radically that other DH conflicts, especially those involving the abandonment of politically important moral oughts, cannot arise. Part of the decision to enter politics is to accept that others might say "if you can't stand the heat, get out of the kitchen". Part of the political vocation is a willingness to stand this 'heat', the responsibilities and hard choices associated with the pursuit and exercise of power. This is not in itself dirtying (the dirt in DC follows from the dispositions and habits cultivated and the approach taken towards morally valuable goals and relationships), but there is an expectation that success might, at some point, require taking on these serious forms of dirt. The dirt that comes from these choices is not in willingly entering a life where, *in extremis*, one might be faced with them, but from the moment of choice itself.

4.2 Moral Luck

ATC-justified dirt often, albeit not always, manifests because of moral luck, especially in cases where the resulting dirt is 'symmetrical' or associated with higher stakes DH-style decisions. Moral luck, as Nelkin (2023: 1) explains, occurs "when an agent can be correctly treated as an object of moral judgment, despite the fact that a significant aspect of what he is assessed for depends on factors beyond his control". Elements essential for the moral evaluation of agents, usually external but also internal, lie outside what the agent could reasonably be expected to control (Williams, 1976: 126). Moral assessment is, at least partially, undertaken from a different position than at the actual moment of choice (Williams, 1976: 131). This includes the extent of an agent's culpability (Nagel, 1976: 146).

Moral luck often dictates both the options for accruing ATC-justified dirt and the ends that obligate such acts. My goal here is not to provide a comprehensive argument defending

the concept of moral luck; such arguments have been made elsewhere (see Nelkin, 2023 for a useful overview). For present purposes, the key takeaway is that denying the role of moral luck, and instead restricting moral assessment solely to those things that fall directly within an agent's control, ends up dramatically reducing, if not eliminating, the scope for moral assessments of individuals (Nagel, 1976: 138). This is because many of a given agent's acts emanate from actions beyond their control. Denying moral luck means that "the area of genuine agency", and with it the scope of "legitimate moral judgement", shrinks into nothingness (Nagel, 1976: 146). Such an argument seems especially apt within the realm of politics, where so much of what agents do and get involved in is beyond their control. Rejecting moral luck therefore restricts the opportunities to discuss dirt, at least the ATC-justified kind.

4.2.1 Types of Moral Luck

There are four ways in which the "natural objects of moral assessment" can be subject to moral luck (Nagel, 1976: 140, named and expanded upon by Nelkin, 2023: 1). Putting aside "causal luck", which is not directly relevant to our questions here, three types of moral luck are important for understanding why politicians may not be able to avoid dirt. Firstly, "circumstantial luck", which concerns the circumstances we find ourselves in and the moral tests that we are asked to face (Nagel, 1976: 145-146). Nagel uses the example of the citizenry of Nazi Germany. They had a chance to behave heroically by opposing Nazism, but also the opportunity to behave deplorably by becoming complicit or culpable in the regime's crimes. Another nation's citizenry might well have failed this test in equal or even greater numbers, but because they did not face this choice, they lack culpability. Dirt arising from counterfactuals and hypotheticals is not the same.

Circumstantial luck, too, can concern things other than actions, such as affiliations. Nagel uses the example of a US citizen in the 1960s who felt ashamed of their country's war in Vietnam and tarred by their unchosen association with it. The anti-war American citizen can be seen as the 'victim' of bad moral luck (Nagel, 1976: 151, n. 9). This is important for politics. A quintessential example of circumstantial moral luck can be seen as DR, where engagement with morally permissible political activities, even in the name of morally important goals, requires engaging with morally corrupting and dirtying practices. The moral character of politicians is affected not only by how they interact with these institutions, but what the institution itself says about the moral character of the politicians within it.

Next, there is "resultant luck", concerning the effects of our actions (Nelkin, 2023: 1). A negligent driver who hits a child is judged as morally worse than one who, by luck, does not; an attempted murderer is seen as less condemnable than a successful one; a gambler who doubles somebody else's investment receives less moral outrage than one who loses it all. One way to unpack this kind of moral luck is to say that it rests on there being, in the positive case, duties that entail success at some outcome rather than merely trying (Gardner, 2004: 76) and, in the negative case, instances where a still-reproachable act of negligence renders our fulfilment of moral duties as dependent upon luck (Gardner, 2004: 63). Resultant moral luck can be something that arises almost out of nowhere. An attempted murderer foiled by a sudden coughing fit (Feinberg, 1970) might have planned the perfect crime, yet they did not commit it. That would-be killer has still committed the morally condemnable act of planning and intending a murder, but they do not have the taint of murder. "In many cases of difficult choice the outcome cannot be foreseen with certainty" (Nagel, 1976: 142), so if a chosen path is particularly vulnerable to moral luck, then that should weigh into our decisions in advance. Subjecting the fulfilment of important responsibilities to exceptional degrees of moral luck can

itself be dirtying: the gambler who gambles with someone else's life chances is a good example of this.

Finally, there is "constitutive luck". This refers to two separate ideas. The first, prevailing, meaning is that this is "luck in who one is, or in the traits and dispositions that one has" (Nelkin, 2023: 1); "your inclinations, capacities, and temperament" (Nagel, 1976: 140). On this view, it is morally correct, for instance, to blame someone for being cowardly, selfish, or prejudiced, even if they have no control over these attributes. Similarly, it is morally correct to blame someone for acting on these dispositions even if they are (currently) impossible for the agent to resist. Agents will of course have moral reasons to try and change these dispositions, but both their internal capacity to do so and whether there is a conducive external environment are matters of luck. The broader demands of morality do not necessarily change based on what an agent is capable of: complying with a broader ethical duty to be charitable or tolerant will be harder if an agent is stingy or distrustful, whether by temperament or upbringing. Complying with duties of charity and tolerance will be harder for such agents, perhaps even impossible, but this does not in itself change what those virtues require (Gardner, 2004: 64). This kind of constitutive luck can also change the morally "optimal" way that different actors can respond to identical situations, in addition to changing how they *will* respond: one agent might find their best course of action is pragmatic compromise, another in the same situation might find the best thing for *them* to do is principled stubbornness. Similarly, what might be good constitutive luck in one set of circumstances will be bad luck in another.

The second meaning of constitutive luck is one implied in how Bernard Williams originally discussed the concept. Williams used the term to refer to "the dispositions of morality" (Williams, 1976: 116), how "intention and motive" are not immune to being conditioned. On the surface this refers to how the basic intentions and core moral principles of specific agents are affected by circumstances. This phrasing also implies a somewhat different

form of luck, one that considers luck in terms of the values and goals we are striving for. If what our intentions are is a matter of luck, then it is also a matter of luck whether we have the *right* intentions. The goals we are fighting for might turn out to be wrong, despite appearing to be right and just and even obligatory to reasonable people in the here and now. Perhaps those goals might even turn out to be evil. There is, of course, a responsibility to examine the worthiness of our goals and if they justify the actions (particularly the unseemly ones) that we are using, yet sometimes, due to conditions beyond our control, we might not be able to determine the right answer.

This might be seen as collapsing into resultant luck, but I would caution against this: resultant luck is about the consequences. This kind of luck is about how we assess consequences, and whether we are using the right moral framework and the right values to do so. For a utilitarian, for instance, resultant luck is about maximisation of utility. They might still be morally unlucky were it to turn out that maximising utility was the wrong principle. Taking the example of Paul, moral luck is evident not only in the sense of resultant luck (if his abandonment of his old self pays off) and the conventional meaning of constitutive luck (if that potential for political success is within him, or whether the broader public is receptive), but also in this other sense too, as whether his political goals truly have the worth he suspects them to have is to some degree a question of moral luck. To avoid confusion, I refer to this second type of constitutive luck as valuational luck. I believe valuational luck to be especially common within politics, where we must be concerned not only with our ability to achieve our goals, but also about the worthiness of those goals. Valuational luck has its limits, of course: it is much less apposite for the TBS than for the agent-relative concerns of a political candidate.

This leaves us with four types of (relevant) moral luck: resultant, circumstantial, constitutive, and valuational. Both the acquisition and justifiability of dirt can be governed by moral luck. This means that agents might face the possibility of justifiably accruing moral dirt,

especially within the realm of politics.[18] It is applied most simply to DR: DR is, in large part, a manifestation of circumstantial luck with reference to the background conditions that one finds oneself operating within, at least at the start of one's career.

4.2.2 Moral Luck and Dirt

If a dirtying act, such as a corrupt deal, does not work out due to poor moral luck, despite being reasonably assessed as ATC-justified at the time, it suddenly moves from being an ATC-justified dirtying decision to merely a morally wrongful act (Nagel, 1976: 142). That it was a reasonable choice is not forgotten altogether. However, the agent took a gamble and morally owns at least some of the consequences. Conversely, an adventurous moral violation can become ATC-justified because of *good* moral luck—a somewhat tricky prospect for DH theorists, as that is not an attitude that should be encouraged, being the mirror of the purist view of leaving the results to fate and fortune. As argued in the next chapter, I object the purist view, and the same goes for its mirror.

Part of the dirt in cases such as rebellion therefore flows from the engagement with moral luck itself. It is dirtying to gamble with one's potential revolutionary allies in this way (Smith, 2018: 205). When rebelling, many of both the deontic *and* consequentialist wrongs, including violence, seem inevitable. Whether those wrongs are ATC-justified is a matter of moral luck. Not only is rebellion liable to moral luck in the sense that the costs of failure can fall upon the unsuccessful revolutionary leader, but whether the required wrongs themselves become ATC-justified can itself depend on attaining success and that success being vindicated by valuational luck.

---

[18] Machiavelli's (1532 [2008]: ch. 25) discussion of the importance of *fortuna* for politicians is especially illustrative here.

Things get a bit murkier with regards to DC. The dirt in DC is less likely to represent bad circumstantial luck and is usually always permissibly attained, regardless of moral luck (though part of the dirt in gaining DC is in accepting a willingness to grapple with moral luck). In contrast the justifiability of both DR and especially DE is dependent on moral luck, both in terms of resultant luck and, particularly if agent-relative and contested goals are at stake, valuational luck (though of course the goals an agent might be striving for via DC can be subject to both types of luck too).

4.3 Many Hands

Politics is a collaborative activity. This can cause dirt and make it tricky to avoid. This will be a recurring theme throughout the thesis: structures, institutions, and political cultures are (often unconsciously) collaboratively designed and maintained, and this can create dirt. Individual decisions are also not exempt from this observation, however. This same interconnectedness of political action means that we are morally affected by, and affect, the decisions of others. Dennis Thompson (1980: 905) used "the problem of many hands" to describe those instances where "it is difficult even in principle to identify who is morally responsible for political outcomes."

The DH analysis, as conventionally phrased by Walzer, is oddly apolitical in the sense that it focuses on lone actors making discrete decisions.[19] In reality, Thompson notes, many such individuals have a role to play in making these decisions (Thompson, 1980: 905). Placing them solely at the feet of an individual actor is mistaken, as is assuming their moral pollution

---

[19] Critiques of this kind would go on to be developed in greater detail by others, including by debating Democratic DH and by looking at staticity, as discussed in the next three chapters. Thompson's identification of Many Hands is just one way in which we can, and should, respond to these problems.

is restricted to individuals. The bureaucrats who implement a decision exercise leeway and discretion in how this policy comes about, which can greatly impact a decision's moral status (Thompson, 1980: 906). Advisors are also highly relevant. They do not actually implement the policies or the decisions for which they advocate.[20] Those with the power of implementation need to take up their advice. However, the possibility that the advisor *might* have no bearing on the outcome does not exempt the advisor from sharing in the moral criticism of it (Thompson, 1980: 911-912).

Thompson emphasises that the difficulties of assigning responsibility should not cause us to abandon individual responsibility altogether. The alternatives are no better, for one. Placing responsibility solely at the top of the hierarchy ignores how those lower down the chain can wield discretion and deserve scrutiny in how they use it, as well as giving those at the top undiluted control over how guilt is processed and wrongdoing is justified to others (Thompson, 1980: 906-907). Assigning responsibility to the collective without regard for structural causes involves blaming the blameless (Thompson, 1980: 907-908). In contrast, personal responsibility allows us to substantiate concepts such as what counts as the least bad alternative, for example (Thompson, 1980: 914-915). Individual responsibility remains normatively important within the context of institutions that disperse such responsibility, including dysfunctional or unjust ones such as DR institutions. In defining personal responsibility, Thompson (1980: 908-909) emphasises both causal responsibility and volitional responsibility, or the possibility that the agent could have done otherwise. This account of personal responsibility is, I contend, compatible with the previous discussions on moral luck and

---

[20] For example, Dana's decision is hers alone: placing it at the feet of an advisor seems to miss the point. However, we can still judge these advisors for the advice that they give. If, say, one advisor knows that Dana *will* follow his advice, that does increase the moral significance of his action.

symmetry: an agent can find themselves in a situation in which every route is dirtying or the ATC-justified decision is dirtying, and yet still meet the standards of volitional responsibility.[21]

Incorporating Many Hands thinking into our DH analysis also creates a much more robust notion of DH more generally: for example, it guards against criticisms that DH analysis rests on an unsustainable 'lone-wolf' view of political conduct (Sutherland, 1995: 485–486; 2000: 223; Shugarman, 2000) by allowing us to distribute dirt across multiple actors. Looking at Many Hands also complements discussions on the democratic authorisation of dirt (Hollis, 1982: 396; Archard, 2013: 777; Levy, 2007; Thompson, 2007; Thalos, 2018: 175-178) and on how dirt can fall upon collective agents (Goodwin, 2009: 237-246).[22]

The TBS is a good case to analyse with Many Hands in mind. From how it is presented in the literature by Walzer and others, it is only Dana's dirtied hands that interest us. Dana's moral responsibility is certainly important, but what about the dirty hands of the torturer? It is rarely the politician who does the dirty work, after all.[23] What about the dirty hands of those who relay the politician's orders, or advise the politician that torture is an option? What about the law enforcement officials who then use the evidence gained through torture? What about the private sector supplier who provided the necessary tools and resources to enable whatever torture methods are chosen? What about the democratic voters who might later reward Dana's decision and later decide they want politicians willing to torture in such circumstances? This dispersion of responsibility can be witnessed in the less dramatic dilemmas of political life too. Those who join Edna's campaign and volunteer for her will plausibly know that they are using

---

[21] Besides, there are limits to how far volitional excuses, such as pleas from ignorance or bureaucratic path dependency, can exempt an agent from responsibility (Thompson, 1980: 912-914).

[22] In this way, it could be used to argue against the notion that Democratic Dirty Hands (or DDH) are incoherent or undesirable. A democratic polity, acting as an agent, can authorise individual agents to incur DH on its behalf or engage in other collective acts that entail collective wrongdoing. I discuss DDH in greater detail later, though not through this lens. Still, it is worth noting.

[23] The notion that torture requires the creation of a "torturing class" (Coady, 2008: 90) will be discussed in greater detail later.

resources derived from a system that both Edna and (presumably) the volunteers themselves consider unjust.

Similarly, take the dilemma facing the political bosses in North Takoma. Legislative aides who are complicit in the gerrymandering process, even if they argued against it at the decision-making stage and implemented it only out of a sense of collective responsibility, all plausibly bear some dirt. Given that all political actors can play some role in shaping a polity's norms and practices, attempts to shift blame solely onto another actor (whether that be a superior, an original author of wrongdoing, or the collective) would ignore their own potential causal role and scope for personal responsibility (Thompson, 1980: 908).

5. Conclusion

This chapter sought to set out the core components of the concept of dirt and thereby address the ambiguities in how it is used within much moral and political philosophy. The VC account that I have defended in this chapter also creates the space for dirt to be ATC-justified. Cases of ATC-justified dirt are therefore understood as arising from value conflicts that, even when the best option is chosen, leave uncancelled moral obligations, generate moral remainders, and ground legitimate moral criticism of the acts and actors involved. This approach not only captures (the vast majority of) the DH cases within the literature, but also a broader range of acts, dispositions, and practices that have double-counted costs that retain non-action-guiding evaluative importance. In the next chapter, I defend the concept of ATC-justified dirt from critiques.

Avoiding this dirt altogether is very hard, if not impossible, for most political actors. This is due to the symmetrical nature of many of the scenarios that present political actors with the option/requirement to obtain ATC-justified dirt, where no dirt-free option presents itself, and the fact that a political career requires running towards scenarios where bad moral luck creates such dirtying predicaments. I do not claim that politicians are powerless to change or minimise dirtying scenarios (as will be recurring throughout the thesis) but doing so requires being realistic and conceptually clear about what sources of dirt can and should be changed in the first place.

Now that the necessary and jointly sufficient features of dirt have been identified, instances of political dirt can be further distinguished based on severity and (as will animate chapters 3 through 6) character. This will allow us to work out when an instance of ATC-justified dirt can be seen as just another part of the cut-and-thrust of politics or as something more regrettable that warrants reform. This gives us the tools to avoid an unimaginative acceptance of the unacceptable at one extreme, or an unworkable utopianism at the other.

CHAPTER II: DEFENDING POLITICAL DIRT

In this chapter, I defend the desirability and utility of the idea of ATC-justified moral dirt that was elucidated in the previous chapter, providing a robust defence of its use by political theorists. There are multiple reasons for believing that ATC-justified dirt is a useful phenomenon when analysing the ethical dilemmas that political agents face and, more pointedly, that viewing these issues through the framework of dirt is better than the alternatives.

I focus on addressing denials of ATC-justified moral dirt. The first is purism, or the idea that certain deontic constraints and rules must be adhered to come-what-may. The second is specificationism, the idea that a seemingly intractable conflict between two values can be solved by respecifying one or more of these values in such a way as to dissolve the conflict. The appropriate response to both views is one of respectful rejection. While accepting these ideas have some normative force and that they bring up useful considerations for moral imagination-style deliberation, I highlight their severe action-guiding limitations and argue that their attempts to recast dilemmas often fail.

I then turn to a more difficult set of oppositional views, consequentialism and threshold deontology. These views both instruct politicians to commit morally distasteful acts when ATC-justified, but hold that, provided they are *actually* ATC-justified, said acts generate no moral remainder and entail no real wrong. ATC-justifiability strips acts of their dirtying status. I think this is a trickier and more important challenge. I respond to this by providing an overarching argument about why dirt is a useful and illuminating concept.

1. Against 'Purism'

The first alternative view I wish to consider is purism: the claim that politicians should simply avoid committing morally dirtying acts. There are two broad types of purism. The first is 'strong purism': politicians should never lower themselves to the base nature of politics. Like all moral agents, they are to be judged wholly on deontological grounds or by the cleanliness of their character. The second type is 'weak purism': that there are certain inviolable deontic constraints restricting how agents can respond to situations or pursue political goods. Purists are absolutists about these rules, though not necessarily about *all* rules.

Both forms of purism have the effect of dissolving some or all moral dilemmas, especially DH ones, by ruling out certain options *simpliciter*. Strong purism also denies that there can ever be a political brand of morality, thus denying DC and DR. Weak purism only denies the scope for DC and DR if the strand of purism in question has a sufficiently expansive list of prohibited actions. Unlike other critiques of DH theorising, purism accepts that moral dirt is real and that it *does* matter. The purist instead argues that the idea of *justifiably* acquiring it is severely misguided. Purist theories thus often provide powerful descriptions of the reality of dirt, even though, as I ultimately argue, their action-guiding content is misguided.

1.1 An Overview of Purism

According to purism, certain acts and practices are incompatible with morality and are never (even ATC-) permissible, regardless of circumstances or consequences. DH dilemmas thus cannot arise. The purist attitude would lead us to forbid any use of torture in the TBS, to advise both Ned and Edna to reject the unsavoury offers even if it dooms their laudable campaigns to

defeat, and to never gerrymander South Takoma. Stronger purists would tell Paul to preserve his virtues and to rise above the gritty demands of politics.

The classic example of strong purism is the Kantian categorical imperative, which stipulates that, when evaluating the moral suitability of any proposed "maxim" or guide for behaviour, we must first ask "whether everyone could act on it, or whether it could be willed as a universal law" (Rohlf, 2020: 5.4). The categorical imperative is an absolute and unconditional rule that binds human behaviour, regardless of circumstances or context, sidestepping the problems of subjectivity and empirical limitations. As the protagonists in all our examples wish there to be "universal laws" against the acts that tempt them, they must therefore not perform them themselves. The categorical imperative allows for no concessions to be made to the evil projects of others, the institutional shortcomings of our political structures, or the cutthroat standards of success that characterise politics.

Deontological moral philosophy centres these rules, morally assessing acts not by their intended or actual effects but by their "conformity" with moral norms (Alexander and Moore, 2021: 2). Consequentialism can still play some role for the purist (especially the weak purist), but only a bounded consequentialism where "deontological reasons [serve] as a kind of side-constraint or override" (Moore, 2018: 374). These constraining moral absolutes are unconditionally true and universal (Waldron, 2011). The deontological form of purism tends to be more outward-facing, seeking to provide moral guidance that is insulated from our own preferences and that avoids reducing others to mere means-to-ends, or lays the groundwork for the implementation of universal laws.

A second form of purism can be inward-facing and focused on integrity of character, a position more akin to virtue ethics (Anscombe, 1958: 16-17). The goal of this form of purism is not to align with and support deontic moral norms, but instead to be the right kind of person. Presumably the right kind of person is someone who does not even *contemplate* torture, refuses

to stoop to corruption and dirty tricks, and embodies their political ideals in terms of both behaviour and favoured policies.

These internal accounts often take a theological tone and content (Griffin, 1989: 36-38), such as a desire to protect the eternal soul (Lovell, 2007: 191). Efficacy is even less of a concern for inward-facing purists. Yoder (1964), for example, argues for an ethic of pacifism and non-violent resistance, solely to display fidelity to Christian ideals that he regards as crucial. If it is efficacious, this is a happy side-effect only (Yoder, 1985; Griffin, 1989: 56). The moral rules are obeyed: the "profit" that comes from abstaining from injustice is "[left] to God to determine" (Anscombe, 1958: 19).

As with DH theorists, purists accept that important moral norms and rules remain even when costly or counterproductive to follow. They are always evaluatively significant, and their violation always involves a discrete act of wrongdoing. However, purists deny that there can ever be cases where it is permissible to break these rules. The moral guidance is simple: obey the rules. This is not to say purists necessarily argue that there is only ever one correct course of action in any given situation. For example, Ramsey (1961: 4) says we can compare courses of action on consequentialist terms, but only those options that do not "directly violate" essential moral commandments. A purist can still choose between "the good or the best or the lesser evil" option by referring to consequences, but only once we rule out any action that is "intrinsically and from the beginning wrong in itself." Dirtying options, however, are off limits.

Purists reject the notion that either the intrinsic or the empirical attributes of politics allow any deviation from more conventional moral proclamations. There can be no separate political morality, and political expediency and empirical calculations are no excuse for moral leniency. Features of politics, such as "moral isolation" and an apparent need to compromise, do not justify us retaliating with deceit and lies of our own (Primoratz, 2007: xi-xii). Yoder (1984: 162) argues that the creation of an independent political morality serves to leave existing

power imbalances and the self-serving assumptions of existing political actors unscrutinised. Many purists view these moral constraints as even more important for politicians than for ordinary people: as their decisions affect so many, we should be even less tolerant of moral shortcomings (Primoratz, 2007: xii).

1.2 The Strengths of Purism

Once it has been established that a moral principle is absolute or that a certain act is intrinsically wrong, this provides all the justification needed for purists. We do what is right because it is right, not for convenience or from on-the-spot consequentialist calculations. Purism does have practical benefits too, which might make it attractive from a political standpoint.[24] It allows us to approach moral questions with clarity and predictability, without needing to worry about strategic errors or empirical limitations (Griffin, 1989: 37). Its action-guiding content is clear and unambiguous. It protects moral judgements, evaluations, and prescriptions from both self-interest and transient subjective priorities (Johnson and Cureton, 2017: 5.4).

There is *some* phenomenological and rhetorical truth to purism, too, since it provides intuitive condemnations of immoral acts. By preventing the creation of a specific 'political' subset of morality, purism prevents politicians from codifying unjust acts as necessary evils or from developing a unique political morality that is prone to self-interest and special pleading. It prevents the corruption that can come from even considering certain vile acts as options (Anscombe, 1958). It is also patently true that, were everyone to follow purist doctrines, many of the difficulties associated with it would fall away. Were everyone to follow Kant's

---

[24] However, I think most of these benefits can be achieved, for the most part and with lower costs, by embracing the idea of dirt. This will be elaborated upon later in the chapter.

categorical imperatives, there would be no problems of moral isolation to confront.[25] This is all the truer for politicians. In this way, purism can be said to both acknowledge the true weight of politics and explain why it should be governed by moral norms.

1.3 Weaknesses of Purism

Ultimately, however, the position we should take with purism is one of "respectful rejection" (as Coady, 2018: 7, characterises Walzer's position). We can reject purism without becoming antinomians, by still accepting certain moral constraints are often practically insurmountable and never wholly defeasible. The purist proclamation is to "let justice be done though the heavens fall" — *fiat justitia ruat caelum*. DH theorists reject this as implausible (Walzer, 2006: 7), ignoring what it would actually mean for the metaphorical heavens to fall. For example, Walzer (1977: 251-252) claims that "supreme emergencies" uniquely justify the breaking of moral rules. A supreme emergency coming to pass would be quite horrific indeed and certainly not in the long-term interest of advancing justice. Sandsmark (2022: 752, n. 25) takes the examples of killing one person to prevent a "nuclear holocaust" or telling a lie to "save a billion people from eternal damnation". The strong purist about murder or lying must say no. This is a "moral absurdity", Sandsmark contends, and it is ultimately hard to view the strong purist in these cases as "[inhabiting] the same moral universe" as the vast majority of us. This is to say nothing of cases, such as the TBS, falling below the level of supreme emergency.

The supreme emergency argument is an obvious challenge for purists. The consistent purist responds by biting the bullet regarding the prohibitions they regard as absolute. A

---

[25] This is one advantage of purism that the idea of dirt cannot necessarily replicate. However, purists lack a coherent strategy to enforce such categorical imperatives and, by distinguishing between different sources of dirt as I do in this thesis, I believe that the dirt approach can deliver *some* victories here.

philosophical interlocutor presented Waldron (2011: 5-6), an absolutist against torture, with a series of escalating catastrophes that could be stopped by torture. Waldron repeated his claim that "torture is always wrong, even in this case" until the other philosopher gave up, looking at Waldron as if he were "mad", an "irrational fanatic". While I neither think Waldron is mad nor think we exist in separate moral universes, I agree with the other philosopher that a world with an instance of torture can be preferable to a world with innumerable casualties. DH theorists reject the proclamation of "let justice be done though the heavens fall." Instead, they proclaim "commit injustice to stop the heavens falling, but do not pretend that it is no longer an injustice." It is from this position, critical of the universality of their action-guiding purpose yet conceding the evaluative absolutism of many rules and constraints, that I think purism should be viewed. This creates what can be called a weak absolutism (Kramer, 2014).

Of the five examples running through this thesis, the moral absurdity claim seems to apply only in the TBS (and even this is debatable). Perhaps we can mimic Waldron's interlocutor's moves and add more and more extreme consequences to these examples. Millions will die if the terrorist is not tortured. If Ned and Edna do not win, there will be a mass humanitarian disaster. If North Takoma is not gerrymandered, the Nazis will win. If Paul does not enter elected office, he will be missing from a crucial parliamentary debate where his insights will prevent the world being blown up. Such a move is unsatisfying. The critic of purism *can* reach for this kind of response but need not do so immediately.

That these acts are justified *in extremis* is not the purpose of those examples. It is instead that they are required *by political considerations*. Here the TBS departs starkly from the other cases. The moral absurdities of purism in the TBS can be solved by adopting the action-guiding proclamations of threshold deontology (to be discussed). I want to argue for a stronger claim: that correct political action, especially in flawed political arrangements but even in healthy ones, requires ATC-justified moral dirt. This goes beyond the concessions made by threshold

deontologists, and delivers, without the unattractive costs, the evaluative benefits of absolutism.

Purism, simply put, is unsuitable for politics. The stronger the purism, the more unsuitable it is. The weaker and less encompassing brands of purism will be less unsuitable, and perhaps their absolutist constraints will never prove a problem. Weaker purists will often prove to be more desirable than political actors *looking* for opportunities to dirty their hands. However, when faced with the sort of examples that I have sketched out, purism seems to entail a weak or irrelevant response. A purist must allow that Ned, Edna, and Paul will all lose their critical elections or be ineffective in office and that the gerrymandering of South Takoma will go unmitigated.

Even though it is right to acknowledge the dirt that comes from breaking these rules, sometimes they should be broken. Politics is a morally valuable endeavour. A sufficiently strong purism places unassailable constraints on the pursuit of political goals, thereby neglecting their importance and leaving political morality as a substantially neutered "morality-only-when-justified" (Cullity, 2007: 61-65). Mendus (1988: 333-334) notes that purism invites us to retreat from the risky game of consequences to comforting moral certainties, a serious limitation when encountered with situations of moral luck. One purist defence is that the blame for bad consequences falls solely on others, yet even an absolutist such as Waldron (2011: 19-21) accepts the unattractiveness and insufficiency of this kind of defence.

More generally, political conduct is inappropriately judged by recourse to abstract moral rules alone. According to Philp (2007: 74), "our judgments [must be] intelligible at some basic level to someone caught up in the particular political context we are examining". It is "antipolitical" to focus solely on intention (Philp, 2007: 82-83) and, similarly, it is antipolitical to focus solely on the purity of individual behaviour. Politicians must act with "a recognition

of the limits of what any individual can achieve" (ibid) and must acknowledge that many of their principles and commitments are made authoritative for others *via* the political realm.

The unsuitability of purism extends beyond cases where the metaphorical heavens might fall. Politics requires a ruthlessness and contextualisation that, at best, strains purist norms. Purism does not just restrict our ability to prevent the heavens falling. It hampers our ability to improve the world in the interim. The unsuitability of purism for politics is highlighted especially ably by Shotwell (2016): a focus on personal purity is wholly unsuitable for shared projects and collaborative efforts, and results in a dramatically limited scope for what individual agents can consider and explore. It also renders nigh-impossible effective coalitional politics.

Taken to its natural conclusion, purism can be seen as advocating an exit or abstention from politics altogether. Those who value their moral purity above all else should not take up a political career, or at least enter it with reduced expectations. I agree with Walzer that those who cannot stand the heat should "get out of the kitchen" (Walzer, 1973: 165). However, the necessary conditions for a successful political career should not be seen as inevitably 'off-limits' for good persons. We want those with healthy and contextualised scruples to take the plunge into politics. Strong purism does not allow this, however. If those of "private virtue" resist political participation, politics is left, by definition, to those who are immoral (Rasmussen, 1972: 66). This exacerbates the problems of extrication, moral isolation, and unsavoury compromises for those within politics.

If we consider politics important — due to its fundamental impact on so many we should — we cannot be content to leave it in the hands of the immoral. The reasons against a purist entering politics will be greater the worse the political situation is. The more we need good moral actors to enter politics, the more reasons there will be for good people to avoid politics. And yet, would this still allow the purist to keep their hands clean? There are reasons

to be sceptical of this conclusion. For example, it relies upon the primacy of 'private' over 'public' morality; the idea that the internal morality of moral righteousness is somehow true morality and the only proper path to a morally coherent identity (Sorell, 2015). Why private morality should be granted this priority is unclear, especially given the moral importance attached to political goals. It is even more unclear why other actors within a polity should give such importance to another agent's private reasons as justification for action, reasons that they may not share, especially when they feel the impacts of public reasons all the more. While not every decision to engage in dirtying behaviour *can* be justified by pointing to wholly uncontroversial public reasons acceptable to "all reasonable citizens" (Eberle, 2021: 481), they should still be enterable into public discourse. Many purist accounts, particularly the more insular kinds, are unable to be discussed in such a way.

Similarly, when faced with what appears to be a DH choice, the purist assumes it is "asymmetrical" (Nick, 2019a), in that abiding by deontic constraints allows for clean hands. Equally, resignation from a position of responsibility does not necessarily grant us clean hands (Hollis, 1982: 393-394). Public officials, via the "many hands" problem, can also be morally implicated in the actions of their broader institutions or the broader collective (Thompson, 1980). Purism might not merely be undesirable. It might be unattainable. Maybe purism, then, is something that can be achieved only by luck. It can thus in turn be jeopardised by bad moral luck (Gardner, 2004: 66; 86).

2. Against 'Specificationism'

The next alternative account I wish to consider here is specificationism (Thomson, 1976). Specificationists argue that, when two obligations seem to conflict, we resolve the conflict not by weighing one against the other, but "by creating an exception to the content of the losing

obligation, an exception which has the content of the prevailing obligation" (Moore, 2018: 376). Essentially, when two obligations clash, we change at least one of them so that the obligations no longer conflict. At its extremes, this is a wholly "adventitious" view of moral conflict, where the appearance of moral conflict serves to "disembarrass" us "of a mistaken view which for a while confused the situation" (Williams, 1973a: 176).

Take Ned or Edna's situation. The only way to win a morally important election is to do the dirtying act. There are two conflicting moral oughts here. The first is "win this morally significant election". The second is "do not engage in this kind of corrupt deal". Assuming that these empirical calculations are as rock-solid as can be, a specificationist would then say that we have misunderstood the moral imperatives. Perhaps the election is not as important as originally supposed, which seems plausible: election victories are primarily important as means to an end, rather than as ends in themselves. Perhaps the first duty is really "win this morally significant election *using whatever means are democratically acceptable*". Perhaps the second duty should be revised so that the wrongdoing of corruption in this context is wholly defeasible and *prima facie*.

Ultimately, I agree with Moore (2018: 377) that while "some degree of specificationism is called for", as it asks crucial questions of those actors who claim to perform acts of justified wrongdoing, it cannot answer all conflicts. In part, this is because to adapt our principles to remove all conflicts between obligations requires adjustments "as rich as the highly nuanced moral world itself" (ibid). Universal specificationism would leave moral obligations constantly underdefined and constantly in flux, rather than providing any real stable guidance. Specificationism also struggles with the "Compensation Objection", which asks why there are still residual moral duties that remain in the event of a justified breach of a right or commitment (Liberto, 2014: 177-189). If the right has an exception carved into it, it is unclear what legitimate complaint those facing a rights violation have.

While specificationism *can* be appropriate, ultimately not all moral conflicts can be dissolved "without remainder" via restructuring one's moral outlook to deny either or both oughts (Williams, 1973a: 179). Specificationist attitudes can be useful for explaining some difficult moral situations, but so too can double-counted impossible oughts and non-action-guiding evaluations. They can exist within the same, admittedly messy, moral ecosystem. It might be possible to create a system by which we can tell when specificationism is appropriate and when it is not: for example, if a right exists to protect *S*, and *S* is best advanced in a particular case by suspending that right, then that suspension might plausibly be built into the right itself (Liberto, 2014: 194-201). If that suspension is needed to promote a countervailing duty to promote another value, it would be mistaken to build it into the structure of the right.

A similar approach to specificationism is Dworkinian-style value-monism, where values such as equality and liberty can be made to "hang together in the right way" (Dworkin, 2001: 90). Much of the motivation for this view stems from holding that these values mutually reinforce one another and have deep and inexorable interconnections (Dworkin, 2001: 82). I think one can accept Dworkin's characterisation of value without parting from the idea of moral remainders. Even if *values* can be made to hang together in this way conceptually, this does not then guarantee specificationism. In the real world, these values can continue to conflict. The immoral actions of others and flawed institutional designs can cause contingent conflicts. Crucially, contingent conflicts give us no reason to dissolve away the dirt. Recognising that a value has been contingently let down best displays respect to that value. If the two conflicting values are conceptually connected, this gives us extra reason to be attentive to the moral remainder: imagine a DH choice between equality and liberty, one where we ultimately choose equality. Not only have we failed to live up to the ideal of liberty, but if equality and liberty truly *are* interconnected, we have unavoidably let down equality too, albeit to a lesser degree.

Moral constraints under a blanket specificationism are too prone to exception and special pleading to really serve as a useful locus for moral discussion and reform. Specificationism also encourages potentially deleterious attitudes in political agents. They could leave this supposed moral conflict feeling that they had been "unprepared" before and having now recognised that they understood one of their moral oughts incorrectly, they will not face a value-conflict of this kind again (Williams, 1973a: 176-177). Sometimes this might be the appropriate takeaway, say if the agent was guilty of naivete or oversimplification.

However, in other cases specificationism is wrong and the standalone importance of the two competing oughts should be reinforced. An agent facing a moral conflict might resolve that they will never be forced into such a position again: for specificationists, they can satisfy this requirement by reimagining their moral commitments, such as by "reducing desire to a minimum" (Williams, 1973a: 177). Ned and Edna, if they were hard-line specificationists, could therefore be convinced to abandon electoral politics altogether as one sure-fire way to do this.[26] This is undesirable if we attach moral importance to both sides of the equation. However, rejecting any role for specificationism creates mirrored versions of these problems. If Ned and Edna take as a complete given their moral assumptions, for example an unconditional belief in the importance of their victory, this is also undesirable. If they feel no conflict whatsoever, we would think less of Ned and Edna: clearly they did not truly care about their proclaimed principles.

In short, specificationism may be appropriate in some cases but it must be conducted with a recognition that there are some genuine value conflicts. The two sides of a moral conflict

---

[26] Thus a risk similar to purism, of good people being left with no morally acceptable way to operate within politics, emerges.

can each carry genuine and non-defeasible moral weight. Sometimes, there might only be a false appearance of a dilemma. Morally valuable ends and constraints *can* clash, however, and a specificationist attitude that obscures this or unduly downplays this possibility should be rejected. On a practical note, the endless second-guessing of one's goals that specificationism demands will likely severely undermine effectiveness in the political arena.[27]

With the electioneering and politicking examples discussed in the introductory chapter, the dangers of both too much and too little specificationism should hopefully be clear. Are these political campaigns really so important as to justify the costs? For Paul, it is arguably a personal decision about whether to give up his innocence, albeit one that has broader implications for how he is viewed and treated. Ned and Edna, however, need to assess whether the political ends are worth the very real wrongs they are committing. This is the same for the gerrymanderers-to-be in North Takoma. The desirable end of their victory, after all, will be somewhat contaminated by the means used to reach it.

The application, or not, of specificationism is most interesting in the TBS. At first, it seems wholly unsuitable. The specific ends, of torture being bad and letting innocent people be killed, seem very strong and not the sort of moral norm we want to drastically weaken. However, specificationism still has a role in asking why the action-guiding threshold for torture is at *n* and not, say, *n*+100 or *n*-100. More interesting still is the specificationist claim that our being greeted with an apparent moral conflict is a sign that we were unprepared. This might be true when this unpreparedness takes the form of previous choices that the agent has made, but is of little action-guiding value in the moment, especially when the lives of others are at stake. If this unpreparedness instead takes the form of improperly considering our moral obligations, however, then this creates its own problems. If an agent believes "torture is bad except in *Y*

---

[27] Perhaps if the dirt is especially egregious or severe, this might be a morally worthwhile cost. If, however, a specificationist attitude is constantly causing a political agent to second-guess grimy acts and the costs of politics as a DC, this second-guessing is more directly unsuitable.

cases", and they believe *Y* is a distinct even if small possibility, then they might create or maintain the apparatuses and institutions of torture even before *Y* materialises. This could lead to damaging degrees of corruption.

In setting out the supreme emergency view, Walzer (1977) holds that the rules of war can be broken only in supreme emergencies. To be sure, this is not a specificationist position, as the wrongful act retains its wrongful character. However, he cites a useful example for exploring specificationism. Walzer argues that a nuclear attack on civilian populations is the sort of move that only supreme emergencies can justify. The atomic bombings of Hiroshima and Nagasaki (and I make no comment on the example itself) were, Walzer (1977: 265-267) concedes, the morally "best" way of securing an *unconditional* surrender from Japan. However, an unconditional surrender was not an immutable or absolute requirement, nor was it required to avert a supreme emergency. On balance, the bombings were unjustified because a less comprehensive surrender could have been obtained without such great moral costs. Walzer thus endorses a specificationist attitude towards the terms of surrender. The terms of a conditional surrender might have been dirtying in this case, however. A specificationist attitude, then, does not necessarily dissolve all dirt, though it can help preserve the most serious deontic constraints.

3. The Challenges of Consequentialism and Threshold Deontology

Purist and specificationist approaches deny DH dilemmas by saying in their own ways that we should keep our hands clean. DH theorists should respond with respectful rejection. Purist attitudes often convey a serious evaluative truth, while specificationist theories pose essential questions for those who claim to be met with insoluble dilemmas. Both views, in their own way, help provide a useful counterbalance to the temptation of accepting the grim (supposed)

realities of politics as fixed and unchangeable. Next, I address two theories, consequentialism and threshold deontology, that endorse the idea that correct political action can rightly require violating supposed deontic constraints, forsaking treasured values, and performing "dirty" acts. The key point of difference from DH theorising is that these views reject the notion that an ATC-justified act can be dirtying. When it is justified to break a moral rule (if such rules do exist), there is no moral remainder.

Often consequentialists and threshold deontologists concede that it is right and useful that the responsible agent *feel* guilty, but they assert that the DH agent is not actually guilty (Nielsen, 2007: 21). They treat these feelings as a useful fiction: useful for character formation, ensuring beneficial remedial work, and ensuring that dramatic methods are not overused. The agent should be treated equivalently to how Williams (1976: 124) proposes we treat a blameless lorry-driver who was behind the wheel in a fatal collision: others will "rightly" try and remove the driver's guilt, "but it is important that this is seen as something that should need to be done." For Williams, however, this sympathy with the driver also coexists with the notion "that there is something special about his relation to this happening, something which cannot merely be eliminated by the consideration that it was not his fault." A moral remainder, in other words. The accounts in this section deny this moral remainder.

3.1 Consequentialism

Consequentialism is the most influential way of denying this moral remainder. Under consequentialism, actions are morally judged based on the outcomes they (are reasonably expected to) bring about. Within consequentialism, the most influential approach, arguably, is

utilitarianism.[28] For utilitarians, the right action is a matter of calculating which option maximises utility. Following moral rules that impede the maximisation of utility in these cases is morally wrong or, at the very least, misguided. Though the costs and harms inflicted will be part of this calculation, no wrong is committed if they are a necessary part of a better whole.

That utilitarianism provides an important foil for dirty hands dilemmas has long been acknowledged. Citing Brandt (1972) and Hare (1972), Walzer (1973: 160-161) originally treated utilitarianism as both an opponent of DH theorising *and* as one of the choices within a quintessential DH dilemma. Walzer often says utilitarian *prescriptions* are correct for politicians. However, Walzer is not embracing the logic of utilitarianism altogether. Unlike utilitarians, Walzer argues that the cost of sacrificing deontological commitments in these instances is real. Utilitarians deny this moral cost altogether. Deontic constraints are, for them, only instrumentally useful, serving as rules of thumb that can be broken without anguish.

There are other forms of consequentialism too. A particularly apposite one is provided by Nielsen (2007: 31-32), who proposes a non-utilitarian consequentialism: we have no duty to maximise utility, but we do have a duty to minimise evil. When "evil is inescapable", and "we have good grounds" for believing that a normally morally impermissible action is the best way to minimise evil, we should perform this lesser evil (Nielsen, 2007: 35). Abstention is not an option, even though these acts remain prohibited for utility-maximisation (Nielsen, 2007: 25-26). We may rightfully *feel* guilt for our role in the lesser evil, but we do not commit a wrong (Nielsen, 2007: 21).

---

[28] Utilitarianism is of course a very complex and diverse tradition, but I bracket most of those internal debates for now.

Threshold deontologists (Moore, 1997; 2018; Arneson, 2018; Cole, 2018; Johnson, 2019) contend that moral rules, when they no longer hold, inflict no dirt once broken. Threshold deontology (hereafter TD) is much more selective about when rules are justifiably broken, with the justifiable (and therefore dirt-free) annulment of rules only happening in highly limited circumstances. TD "holds that deontological norms govern up to a point despite adverse consequences; but when the consequences become so dire that they cross the stipulated threshold, consequentialism takes over" (Alexander and Moore, 2021: 4).[29]

This approach bears similarities to those taken by Ross (1930) and Childress (1982), according to which an otherwise inviolable *prima facie* value can be violated without remainder if a more pressing *prima facie* duty demands it. Where TD differs is that it is far more selective about when these deontic constraints are waived: a threshold needs to be passed before a rule can be waived, rather than a specific choice between two values causing such suspension.[30] Before this threshold, consequences can matter (Moore, 1997: 723), but not to the degree that they overrule deontological commitments in either a prescriptive or evaluative sense. Both DH and TD, in essence, want to deny action-guiding moral absolutism without denying the importance of rules themselves. They disagree on what status these waived rules should hold.

3.3 Issues with Consequentialism

---

[29] Here I use sliding-scale/ratio TD, where instead of there being a single "fixed threshold of awfulness" at which all deontological constraints are upturned, the threshold depends on the deontic constraint in question. This is the stronger version of TD.

[30] TD also bears an action-guiding resemblance to the "supreme emergency" view Walzer (1977: 251-267) ultimately adopted, where certain extreme threats justify abandoning our regular deontic constraints. The supreme emergency view does not deny the reality of resulting moral dirt.

My main argument against these views is that dirt is real, it is useful, and that it matters. I consider this the key argument for rebutting both consequentialism and TD, leaving DH as the strongest opposition to purism and specificationism. First, however, I want to raise some specific objections to each account, starting with three criticisms of consequentialist views of political decision-making. The first is that the 'ends' of politics, which are the consequentialist's currency for determining rightness and wrongness, are not in themselves predetermined. When considering whether to break a deontic constraint, it is best not to leave unexamined the goods and normative assumptions that are motivating us to do so (Coady, 1990: 265-267). This can lead to us neglecting how the ends can be polluted by the means (Shugarman, 2000: 238) or having too stringent and inflexible a definition of necessity (Walzer, 1977: 265-267). Is Ned's victory truly necessary, and is the moral value of his victory tainted if he takes the deal? There is a *limited* role for specificationism that consequentialism obscures.

The second is that the assessment of political actors cannot solely be consequentialist, no more than it can solely be deontological. Political actors cannot be judged solely by consequentialist concerns: they must also be judged by, and show a responsiveness to, the values, claims, and commitments that are embodied in those decisions, even if this is not action-guiding (Philp, 2012: 638-642). We judge politicians by how well they live up to their values and how well they advance their values in the real world. Similarly, there are reasons against expecting them to live completely detached from the contemporary moral rules that bind the rest of us. A "beneficent audacity" (a term applied to DH by Coady, 2018: 1) ignores the social functions of rules and allows politicians to self-interestedly set the standards by which they guide their behaviour.

The third point is that consequentialism neglects "agent-relative" reasons (Nagel, 1986: 152, discussed further by Goodwin, 2009: 39-40 and Yemini, 2014: 170). Agents have personal projects and commitments flowing from "reasons of autonomy", as well as special relationships

such as family and friendship ties. Consequentialist analysis gives these reasons only instrumental significance. Usually these reasons, unless they represent specific agent-relative articulations of broader mandatory principles, do not generate extreme dirt when broken nor do they tend to justify dirtying behaviour. They do, however, frequently condition what justifies the agent's duties to pursue agent-neutral reasons. If Paul, for example, has certain especially demanding agent-relative obligations to his family, that will condition how he should allow politics to change his temperaments or, conversely, how dirtying these changes truly are for him.

In all five examples from the introductory chapter, consequentialism will recommend something like what the DH theorist recommends yet denies that there is any resulting moral remainder. In cases such as the TBS, this neglects the true moral gravity of authorising torture. In the case of Ned, and by extension with Edna and Paul, what is important for Walzer is that we do not just want *any* politician to dirty their hands by accepting these deals. We only want the 'right' politicians to make these deals; those with the correct political projects and attributes. For Walzer (1973: 166), a key part of identifying these politicians is the fact that they display the appropriate scruples in both their attitudes and attributes. These scruples lose their moral importance under consequentialism, becoming at best a useful fiction and at worst a pathology that neglects the social importance we place on them. Finally, in the case of the Takomas, a wholly consequentialist attitude will lead to a corrupting normalisation of gerrymandering, one that is insensitive to the democratic harms that motivate condemning gerrymandering in the first place.

Threshold deontology has two main weaknesses compared to accepting the possibility of ATC-justified dirt. First, it has limited explanatory scope. DH considers cases that TD does not. TD focuses, by definition, only on extreme marginal cases, where a deontic constraint suddenly ceases to hold. For minor acts of wrongdoing, TD theorists can say that these thresholds are much lower, cancellable under much more mundane and quotidian occurrences, such as if you are already ingratiated within the dirty terms of a competition. That is how a threshold deontologist might argue for their theory's applicability in cases such as Edna's or Paul's. I am unsure how this threshold-oriented approach can capture dispositions and habits of the kind that effective political conduct requires, or how meaningful these minor deontic constraints remain if subject to repeated waivers under normal conditions. This is an extrapolation, though. Normally TD accounts are silent on these diurnal matters.

In the usual cases to which TD is applied, deontology is suspended only under certain, usually extreme, conditions, resulting from conflicts between consequentialism and deontology. The idea of dirt that we have discussed thus far rejects both requirements: ATC-justified dirt need not arise within eliminative emergencies, nor need it only concern trade-offs between consequentialism and deontology. TD does not, in its usual iterations, allow for the more quotidian forms of dirt that this reformulation enables.[31] It is true that there are dangers to viewing political dirt as ubiquitous. These include psychological dangers such as desensitisation. However, I also think that denying the moral costs of an ATC-justified action will lead to its own psychological dangers. I shall return to this thought in the "Why Dirt Matters" section.

---

[31] As discussed in the previous paragraph, I do not see the attraction in adapting TD to allow for this.

Secondly, TD problematically elevates the role of thresholds in our core moral reasoning. This results in treating individuals as 'moral ballast'. Individuals only have an aggregative importance: does their presence help reach the moral threshold? Affected parties have no moral role beyond this, and this leads to various counterintuitive implications. There is a broader point here too: in essence, thresholds should not function as an on-off switch for the moral and evaluative power of values and ethical considerations, even if they must serve an action-guiding function.

What is the moral ballast critique? In short, individuals are reduced to nothing more than commensurable inputs into a moral calculation. Individuals are used to determine whether a certain threshold for action has been reached, and then, once that threshold has been breached, "treats those whose sacrifice can save the endangered as expendable resources" (Alexander, 2000: 904). Larry Alexander (2000: 900-905) illustrates this argument with an adapted version of the TBS, the Terrorist's Mother. A terrorist has placed n people in danger and the only way to secure the terrorist's surrender, and save the lives of the innocents, is by torturing his innocent mother. The threshold at which the strict constraint against torture shall be waived is, coincidentally, at n. Thus, following TD, we begin torturing the terrorist's mother without incurring dirt.

However, suppose that the number of hostages drops below n for whatever reason. Maybe one escaped, or was rescued, or was released or even killed by the terrorist. This brings the number of hostages to n-1. Torture is now no longer justified. Though its action-guiding implications are absurd, I do not think this is a *decisive* objection to TD: the case's conditions are similarly very narrow. This example raises more important and pertinent debates, however. One is about the nature of thresholds, the other is about how these thresholds are used in practice.

One defence of TD is the argument that thresholds are endemic to moral philosophy. A common critique of TD is that deciding where to place any threshold will be an inherently arbitrary task (Alexander, 2000: 905-910; Zamir and Medina, 2010: 53). Why at *n*, rather than at *n*-1 or *n*+1? This is an important question. However, designating a threshold seems inevitable for *any* pluralist moral account (Moore, 1989: 332). *Any* moral view that identifies a "precise point", a tipping point for action-guidance, at which actions go from permissible to impermissible affirms these steep discontinuities (Arneson, 2018: 5). If thresholds are an insurmountable problem or a cost not worth sustaining, only monistic theories of value survive. That is a move neither threshold deontologists nor DH theorists want to take. Thresholds are tricky things, but we need them.

I do not think this response completely nullifies the arguments against TD. For other theories, thresholds play a secondary role, as articulations and expressions of other moral truths. For threshold deontologists, they play a more central role. Granted, they follow from the threshold deontologist's prior deontological commitments and fundamental principles (Moore, 2018: 375-377). However, the threshold deontologist (usually) conflates a rule's action-guiding content with its entire moral profile. This amplifies the risks of identifying thresholds. If thresholds are expressions of other moral truths, it will be easier to assess whether these thresholds have been set in the right place. If the thresholds themselves are associated with moral truths, taking a step back to evaluate whether those thresholds are correctly placed becomes trickier. A threshold that is incorrectly placed, due to epistemic limitations or subjective biases or other such reasons, becomes far more damaging.

One way to get around this critique is Arneson's (2018: 9) proposal of "scalar judgement". Even though action guidance necessarily operates at a steep discontinuity, moral

evaluations need not follow. An act can be less or more wrong. Torturing at *n*-1 would only *barely* count as wrong. Torturing at *n*-25, however, would count as a much more serious wrong. I think this is the correct way of incorporating thresholds into our moral philosophy. It allows us to track how an act goes from prohibited, to permissible, to required. However, I think Arneson's response gives up much of what is distinctive about threshold views and ultimately leads us back to the position of DH. Why not express this scalar spectrum from wrongfulness to rightfulness in the language of dirt? This seems to better capture what differentiates an act that is *barely* wrong from one that is overwhelmingly so, and vice versa with the barely right. In short, thresholds have a crucial action-guiding role. They should not then be given the *additional* job of dissolving dirt once this threshold has been passed.

Cole (2018: 18-21) has a second practical response. Cole argues that the Terrorist's Mother example rests upon faulty premises regarding how we use thresholds to guide our actions. We are justified in torturing the mother, Cole argues, only once we reach *n*. However, if we were to begin torturing and *then* fall to *n*-1, that does not mean we cease torturing. Cole draws out a distinction between *starting* torture and the threshold below which we are required to *stop* torture that is already underway. I think this works as a practical solution.[32] It does raise the question of what we do when we reach that second threshold (the language of extrication might be useful here), but it minimises exposure to counterintuitive strange puzzles.

However, when all is said and done, we still would have tortured to save only *n*-1. The gap between the evils has shrunk. The agent is justified in *continuing* to torture, yet not so in torturing in the first place. Continuing an act of torture to save *n*-1 might be justified, but not having initiated that act in the first place. The agent's actions are still, retrospectively, ATC-wrong. A turn of good fortune for the hostage is, counterintuitively, bad *moral* luck for the

---

[32] Furthermore, the amount of physical and psychological harm inflicted between no torture and some torture is plausibly far greater than the differential harms between some torture and some more torture.

agent in charge of their rescue. Cole's point is well-taken as an action-guiding prescription, but is only bolstered by using the idea of dirt to describe its rightful and wrongful elements.

3.4.2 Moral Ballast

The biggest problem for the Terrorist's Mother scenario is not these practical problems, but the moral ballast critique (Alexander, 2000: 901). The terrorist's victims ultimately serve "no purpose other than to create a duty to save someone else, or to override a right that stands as an obstacle to saving someone else." They are building blocks towards triggering the duty to rescue the whole. How does this respect the individual victims as moral ends? They boost others to the level at which these others gain a right to be rescued, ultimately having no relevance beyond determining whether we can intervene to save the nth life.

Suppose the authorities hold the terrorist's two innocent siblings captive. Torturing either one is sufficient to save the lives. Torturing both is an unnecessary evil. We may torture one but not both. Morally, these options are identical. Those who are being sacrificed are thus treated as "expendable resources". The siblings are treated as commensurable with one another without any specific moral claim attached to the individual *qua* individual (Alexander, 2000: 905). In practical terms, the moral ballast problem can be partially mitigated with a mixture of Arneson's and Cole's arguments. These solutions do not address *whose* rights are being violated, however. Instead, a specific obligation should exist to every individual who makes up n, and a specific wrongdoing suffered by every individual who might be chosen for the sacrifice needed to save n.

Moral dirt proves to be a better way of ensuring that we properly appreciate the moral weight of every individual involved. Moral dirt, using Stocker's language, treats the moral claims of all individuals involved as "double-counted" oughts. Even in the face of an

overwhelming action-guiding argument, the rights they possess retain their evaluative force. Using dirt in this way, to double-count the fundamental rights of individuals, best respects the "unique and intrinsic" importance of each individual involved in these deliberations (De Wijze and Goodwin, 2009: 534), and best respects the value of responsibility to persons that is pervasive in politics (Gowans, 1994: 117-154).

4. Why Dirt Matters

Both consequentialism and TD share a far more fundamental deficiency. Recognising the moral dirt that can be attached to ATC-justified decisions is a valuable, compelling, and useful endeavour. Consequentialism and TD both miss out on this.[33] The argument in this section focuses on two points. The first is to demonstrate why the real presence or absence of dirt, as opposed to the mere appearance of dirt, matters *instrumentally*. The second part of this section's argument is that, if we want to take deontic constraints and norms seriously (which I argue, contra consequentialists, we should), then a key part of this is recognising that dirt follows from their violation. Another consequence of explaining why dirt matters is that it also gives political actors reasons why, *ceteris paribus*, they should wish to avoid dirt. These reasons go beyond the political benefits of appearing to have clean hands. Minimising exposure to dirt is useful for both positive and negative reasons.

---

[33] The arguments about why dirt matters can also be used to prevent an overcorrection to the purported overeagerness to identify cases as DH. Coady (2018: 2) is highly critical of describing all morally justified exceptions to norms such as honesty and trade-offs between values as dirtying. I think he is right to draw this distinction, but by accepting the role that dirt can play, we can see how it can be useful to see these choices as sometimes being more complicated than merely a morally permissible exemption.

I have two purposes here. Firstly, to demonstrate that this debate *is* important and the existence of dirt is of more than just theoretical interest. Some pro-TD critics of DH argue that the existence or absence of real moral wrongdoing is ultimately nothing more than a "terminological" dispute (Johnson, 2019: 6). Functions like tragic remorse can be useful for goals such as character formation (a point endorsed by consequentialists such as Hare, 1972), but there is no need, they argue, to identify the dirt as real. Those who say this debate is merely terminological are missing out on something important. Secondly, and more importantly, to show that this debate should be settled on the side of moral dirt arising when certain acts and practices are committed even when ATC-justified. De Wijze and Goodwin (2009: 538) identify four areas where acknowledging dirt adds to standard political decision-making processes. Doing so:

1. Recognises how, under a "complex immorality", there may well be "no morally cost-free course of action". Every (in)action can carry with it moral costs that culpable agents must reckon with.

2. Prevents imposing an inflexible moral template on "cases of intractable moral conflict" that call for more nuanced accounts. Exclusively deontological or consequentialist frameworks, when applied to a specific decision, ignore the costs of neglected consequences and violated principles respectively.[34]

3. Via double-counting, ensures dirtying acts are properly costed rather than treated as simple inputs into cost-benefit analyses. This cultivates "a habit of reluctance", an important psychological barrier to block the "happy acceptance of the intolerable".

---

[34] Threshold deontology, it is worth emphasising, usually neglects consequences. Once the threshold has been breached, however, the violated principles are rendered largely forgotten.

4. More accurately captures the moral phenomenology of real moral conflicts. Accounts that deny the reality of the dirt fail to explain why we should accept such a gulf between the moral truth of such difficult choices and how good actors experience them.

I agree with these arguments. The third is especially important, addressing an important question: once you allow that these norms *can* be broken, how do you prevent them being broken unjustly? Here, TD and DH share a common problem. For TD, Moore (1997: 723) acknowledges the "psychological danger" in acknowledging the existence of thresholds. Likewise, having the language of DH available might lead to "expansion and exploitation" of this scope for transgression and lead agents to accept as given the circumstances and goals that seem to render DH acts necessary (Coady, 2008: 89).

However, I think the psychological danger is likely greater for those who deny the moral reality of ATC-justified dirt. If you get the choice right, no real moral dirt attaches to you. There are costs, but your own moral integrity remains whole. For the DH theorist, dirt attaches to you regardless. There is *no* way to violate important deontic constraints without incurring dirt. If there is any value in encouraging people not to break these important moral rules too easily, reiterating that even the right call will not *completely* vindicate them can engender greater caution. Not acknowledging this dirt is to "trivialize terrible actions and undermine the moral horror of the harms done" (De Wijze, 2009: 309). Acknowledging the dirt raises the bar for undertaking dirtying acts and ensures that they remain a last resort, as agents will know that, even if vindicated by outcomes, their moral violation still warrants an important "layer of self-reflection" (De Wijze, 2012: 199). This is less appropriate language for more quotidian and grimier acts of ATC-justified wrongdoing, but the underlying rationale remains intact.

One area where the importance of recognising the moral dirt is especially prevalent is in discussions of accountability. This section will look at three aspects of this discussion: punishment and forgiveness, responsible public debate, and questions of corruption and exemplarity. In short, the recognition of moral dirt as being real, even in cases where it is justified, provides opportunities for expiation and moral repair; provides a more accurate framework for assigning responsibility in complex cases; and better recognises the importance of mitigating corruption and using the plurality of differing moral characters.

4.2.1 Punishment and Forgiveness

A common debate within the DH literature is the question of punishing DH actors. Walzer (1973: 178) argues that punishment is needed to provide "expiation" and "personal salvation" for the agent, to keep them tethered to the rules that they have parted from, and to provide the politician with an appropriate fusion of both internal and societal accountability. De Wijze (2013) argues that punishment properly acknowledges both the evil suffered and the agent's causal responsibility in said evil, a combination that can fittingly be described as "dirt". Punishment then provides a cathartic function in dealing with the dirt that is attached to this wrongful act.

This view is far from uncontroversial. Both Levy (2007) and Meisels (2008) regard punishing ATC-justified DH acts as both incoherent and immoral. Punishment might discourage people from breaking these moral constraints *when it is the right thing to do*, institutionalising the very moral inflexibility that DH as a concept is meant to prevent (Levy, 2007: 40-42). Punishing a specific actor also denies just how widespread the true responsibility

for DH is (Levy, 2007: 39). Meisels (2008: 157) emphasises the risk of a "slippery-slope effect" surrounding DH practices and acknowledges the "severe moral blemish" that DH acts inflict (Meisels, 2008: 154), yet still condemns the idea of punishing DH actors. If an agent "acted in good faith and with the utmost caution, with risk to his own liberty and career" (Meisels, 2008: 173), punishing said agent is wrong. They should be "legally exempt" even if not "morally vindicated" (Meisels, 2008: 171).[35]

However, Levy and Meisels do not deny the importance of social accountability. Levy emphasises the need for broader societal recognition of the responsibility for DH, beyond the single responsible agent. Meisels still requires proof of good faith and prudence, and allows for non-legal 'punishments' such as political defeat. Neither conclusively refutes the possibility that punishing DH actors might itself be an ATC-justified dirtying act, as Walzer suggested. Other alternatives to punishment include "no-fault forgiveness" (Roadevin, 2019), placing obligations on agents but without blame, and "positive responsibility" (Kellison, 2019: 165-156), which involves accepting a specific duty to mitigate future harms. Nick (2022) discusses the function that apologies might play in the aftermath of a DH decision.

Ultimately, the exact nature of this debate around punishment need not be resolved here. The common ground between Walzer, De Wijze, Levy, Meisels, Roadevin, Kellison, and Nick is that some kind of socially-acknowledged expiation, some remedial duty, is needed. As a shorthand, I will describe this as 'punishment'. In this more expansive form, it is likely to be applicable to all kinds of dirt: with DR, there are reasons to reform political actors too once the corrupting institutions are rectified, and DC actors might need to relearn dispositions if they are to exit politics.

---

[35] If a DH agent is morally unlucky, space for punishment seemingly reopens under this account. This emphasis on outcome renders Meisels's account liable to Walzer's (1973) critique of the Machiavellian view, that this results in a framework where only prudential concerns limit the behaviour of politicians.

The threshold deontologist and the consequentialist might endorse punishment. For consequentialists, this endorsement is contingent based on consequences, of course. For threshold deontologists, they might agree with arguments about catharsis or restoring these deontic constraints (and will likely be joined by rule-utilitarians on the latter point). Nothing in their theory prevents these reasons from being considered or necessitates a justifiably broken threshold being broken for *all* purposes, including legal ones. Yet if the punished agent did nothing wrong, how is punishment, in the sense of incurring penalties or obligations onto the DH agent, justifiable? How is this not itself a case of doing wrong to do right? Consequentialists could bite the bullet and ignore this. Threshold deontologists lack this luxury. They might avoid this by rejecting punishment and its alternatives that entail an admission of wrongdoing, such as no-fault forgiveness, for these reasons. This is, however, to cede that rejecting the moral reality of dirt dramatically limits our options for enforcing these psychological barriers.

4.2.2 Dirty Hands in Democracies

Acknowledging the dirt thus allows the development of a multitude of coherent and attractive accounts about what sort of "expiation" is required. As Levy indicates, these notions of accountability are not only important for individual agents, but for broader society too. As a more malleable currency, dirt can be distributed rather than concentrated on any single individual, unlike the all-or-nothing right-versus-wrong dichotomy surrounding wholly action-guiding theories. This is useful for both Many Hands cases and for broader questions of democratic culpability.

Dirtying acts can be democratically authorised. The broader voting public shares in this dirt, as they entrust politicians with positions of power where DH conflicts regularly occur

(Hollis, 1982: 396; Archard, 2013: 777; Levy, 2007; Thompson, 2007; Thalos, 2018: 175-178; cf. Kellison, 2019, who makes a similar argument regarding the immoralities of war). Broader society's demands also help dictate how much hypocrisy and dishonesty politicians need to exhibit in the aftermath of DH acts (Bellamy, 2010). The language of dirt can encourage responsible public reckoning, making both the society and its politicians assess whether their DH actions were justified and what sort of resulting expiation is required. Members of the wider society can acknowledge their causal responsibility in any evil suffered, even if they were justified in playing such a role.[36]

This emphasis on democratic authorisation and broader societal culpability helps mitigate a common charge against DH: that it cannot coherently arise within democracies. In part, these critiques erroneously assume that DH scenarios inevitably call upon lone-wolf actors to substitute their own moral judgement in place of whatever is prescribed by the relevant democratic norms (Sutherland, 1995: 485-386; 2000: 223; Shugarman, 2000). I reject this denial of Democratic Dirty Hands (following De Wijze, 2018; Nick, 2019b) in large part because the assumption that DH protagonists must be lone-wolf actors untethered from democratic authorisation is flawed.[37] Denying dirt as TD and consequentialism does, however, increases vulnerability to critiques of this kind. By eliminating dirt from the equation and instead thinking about these cases as eliminative emergencies befalling individual actors, the scope for public reckoning and scrutiny is reduced. We lose out on a language that acknowledges the moral wrongs that can substantiate the necessary "formal retroactive

---

[36] We could also draw distinctions between those whose contributions were ATC-justified and those who were not. Edna's contribution to a dirtying campaign finance system, for example, might become ATC-unjustified if she could have, via her own initiative, nullified its unfair advantage or democratically repealed it. If she has no choice for now but to go along with it, her actions are more likely to be justifiable.

[37] This *can* be the case with Dirty Episodes, where the distinct dirtying acts may not be democratically authorised. Even then, though, many more DE cases will be Many Hands cases. The TBS certainly cannot be a lone-wolf case, and though Ned's might be, it is very likely he has campaign advisers, supporters, and party allies who can also be part of this decision. The lone-wolf characterisation is especially ill-suited for Dirty Rules and Dirty Careers discussions, as the dirt in those cases stems from the rules of (democratic) competition that all actors must engage with. In these cases, accruing ATC-justified dirt is made possible by the presence of other democratic actors and responsible agents.

processes" (Nick, 2019b: 72-73) needed to make this kind of morally distasteful act compatible with democracy.

The psychological dangers exacerbated by denying dirt can affect broader polities as well (De Wijze, 2009: 317), such as when political actors are far too ready to commit a dirtying act or take an all-or-nothing attitude towards whether deontic constraints hold in a given context. These risks are amplified further by the partialities that are common within politics; for example, the existence of greater obligations and exclusive duties towards one's compatriots or one's fellow partisans. This is a concern not only in politics, of course, but in any situations where there are special obligations or duties. A parent might be justified in prioritising their child's fortunes over the entitlements of others. Lawyers and advocates are often charged with sticking up for their clients, in the hope that the adversarial system itself reaches the truth. However, these justifications, even if real, do not automatically wash away the moral cost that comes from leaving other moral duties behind. Acknowledging this dirt allows us to be cognizant of the costs, and therefore hopefully the limits, of the moral force of these powerful partialities.

4.2.3 Impacts on Character

Breaking deontic constraints can both affect one's chosen moral character and cause corruption. What is meant by one's 'chosen moral character'? Essentially, the appropriate answers to many political and moral choices will depend on *who* is facing them. Their skills, projects, dispositions, previous decisions, and societal roles can all affect what is the 'right' way to respond. A dirty act may be impossible for one person, yet morally obligatory for another: if Dana cannot see her order through, or the North Takoman legislators lack the technical know-how to gerrymander, their options change. These conflicts do not arise in a

vacuum, and nor do they star featureless protagonists detached from the broader context and their real emotional commitments. This "episodic" structure leaves a "psychologically impoverished, abstract, and apolitical self" at its core (Tillyris, 2019a: 10). Politicians should not be viewed as static choosers, starting from a blank slate at any given moment, "operating in a vacuum" divorced from "any social identity and *telos*" (Tillyris, 2016a: 166).

Take the idea of moral exemplarity as an example. Effective political movements, Dovi (2005: 129) argues, require not only actors who are willing to incur DH but also those who are willing to embody their principles even when doing so is difficult. To ignore this and say that moral exemplars must acquiesce to the logic of DH in every case is to hamper a movement's effectiveness.[38] One does not need to go as far as moral exemplarity, however. In most contexts, it is politically *convenient* to be seen as not violating key moral rules. The politician who breaks the moral constraints against wrongdoings such as torture and dishonesty, then, will face second-order dilemmas about whether to admit to their dirt. To be successful, these initial departures from moral rules might require that additional hypocrisies are committed afterwards (Tillyris, 2016a: 165-166; 2016b). DH politicians are asked to wear "clean gloves", and while this means the politician hypocritically conceals some of their moral wrongs, it also serves as an important constraint on the normalisation or overuse of dirtying behaviours (Bellamy, 2010). These considerations all affect what an actor ought to do.

The key idea here is that different people facing similar choices can relate to them in fundamentally different ways. An exclusive focus on abstract principles and their strict action-guiding content leaves us unable to deal with the question of what sort of person we wish to be (MacIntyre, 2007: 118), and, in politics, what sort of politician we should be (Tillyris, 2016a: 169). There is also an *a posteriori* element to this (Mendus, 2009: 118-120). Individuals are

---

[38] That the exemplar's choice of exemplarity serves as part of a broader effective social movement, and thus their exemplarity cannot be said to tell against them morally by virtue of others 'picking up the slack', is admittedly to some degree a question of moral luck. The same goes for those who completely eschew exemplarity in favour of pragmatism.

not appointed to roles of political power in a vacuum: they are usually appointed to those roles *as the individuals that they were before*. There was a reason, in most cases, why *that* person was selected for the job. This pluralism of perspectives and standpoints can be a strength within political movements and institutions. Different people bring different skills and insights to the table, and as different people, the correct way for each agent to engage with the dirt of politics may vary quite dramatically. There will be moments, of course, where the action-guiding moral obligations become the same for everyone. However, though the *action-guiding* content of morality can become the same for everyone, the evaluational content of morality need not be. The costs are different. The impact on moral character varies.

Integrity is a strong example of this. What counts as integrity for one agent will not count as integrity for another. Integrity is about "standing for something" (Calhoun, 1995). That we adhere to our integrity is of interest to our "co-deliberators" as we seek the best possible answer to shared moral questions (Calhoun, 1995: 258). Integrity, as well as being an important individual virtue, is socially important in proving the agent can be relied upon as a principled actor who stands by their commitments (Mendus, 2009: 26-27). Integrity can also provide useful ways for society to handle questions of moral luck (Walker, 2003), while Payson (2017) argues that, in cases of structural injustice, integrity can help keep notions of social justice alive when we must engage with these same unjust structures.[39] Though it is important that broader morality serves as a check on individual integrity, at times integrity can win the day in terms of action-guiding power (Mendus, 2009: 32-33): the partial commitments that constitute integrity can themselves be intrinsically morally important.

How does this all relate to dirt? Dirt provides the way to differentiate between these differing relationships to one's role. It does this both before and after DH conflicts. Before, in

---

[39] Payson calls for an idea of meta-integrity to be followed in cases of structural injustice. This is very appropriate for many of the cases we look at here, including DR, DC, and what Corvino (2015) calls Structural Dirty Hands. It will be discussed more in Chapter 6.

the sense that the dirt (or lack thereof) that an agent has can affect their moral character *going into* a predicament of this kind, and thus the appropriate response, such as, for example, if they have a pre-existing claim to moral exemplarity. What DC has an agent signed up to, for instance? After, in the sense that the dirt that follows a DH choice will affect the agent's disposition and moral attributes looking forward, such as whether they will need to turn to dirtying tools like hypocrisy or secrecy to maintain their reputation. Dirt thus makes legible the impacts on integrity, as well as explaining their moral and political coherence. Acknowledging the moral remainder thus creates a better sense both of what the right answer is in moral conflicts and what the costs of that answer truly are.

Chief among these costs is corruption. Corruption is the degradation and erosion of the "morally appropriate operation" of both agents and groups (Coady, 2008: 94), as well as institutions (Miller, 2007). Intentional or culpable corruption is a *prima facie* immorality. Much like loss of integrity, individual corruption is not only a problem for the individual, but also a problem for others. Corruption entails the creation of "conditional intentions" to perform wrongful actions (Coady, 2008: 94-98) and the cultivation of dispositions deeply unsuited for ordinary morality (Curzer, 2006: 47-50).

The corruption of institutions and groups can consist of developing the necessary institutional capacity and expertise to carry out immoral tasks or to spot potentials for their use. Practices such as torture are "not for amateurs" (Shue, 2006: 236) and require a capability to be developed—a "torturing class", in other words (Coady, 2008: 90). Once this capability has been developed, it is hard to stop its "metastatic tendency" to be used as a "shortcut" in other areas (Shue, 1978: 141-143). An "economically rational" use of such cruelties is a "psychological fantasy", ignoring how such attributes escalate (Shklar, 1984: 213). Institutional corruption can also take the form of the problem of bad choices (Dovi, 2005: 132-136), where the "scope of policy debates" are restricted as morally preferable courses of action

lose their political feasibility. Institutional corruption can also mean the fatal undercutting of important deontic constraints: for example, if compliance to rules such as the laws of just war, which are meant to constrain considerations such as necessity, instead becomes dependent on these very considerations (Bellamy, 2004: 838-839). As in "noble-cause corruption", corruption itself can be the dirtying action that is ATC-justified within a DH scenario (Miller, 2007).

Acknowledging the dirt that comes from corruption better allows us, as both individuals and members of a broader group of society, to reckon with and process the costs of dirtying ATC-justified decisions. To view this corruption as a cost, but to deny that it involves any dirt because it was part of a justified whole, seems contradictory and counterproductive for character formation. Instead, corruption should always be viewed as a moral cost. Even if a given increase in corruption was part of a justified whole and therefore a cost worth paying, this does not cancel out the additional moral strain it adds to society or the corrupt character of the resulting situation. Threshold deontologists and (many) consequentialists will view corruption as, *ceteris paribus*, a wrong. That its source was originally part of an ATC-justified whole should not cancel this out.

4.3 The Usefulness of Deontic Constraints

The paradigmatic DH cases involve deontic constraints (temporarily) losing their action-guiding content. Certain deontic constraints, I argue here, are more than mere contingencies or instrumental guidelines. They have intrinsic value, and thus dirt arises when they are broken. They should thus be repaired.

As Walzer (1973: 169-172) argues, rules serve important social functions as focal points for ethical discussion and reform. It is not unreasonable to expect violations of these

rules to be justified and defended.[40] Those who ignore rules completely neglect to take these functions seriously. Some utilitarians may respond by citing the social function of rules as a reason for *instrumentally* overvaluing them, akin to Hare, but still contend that when justifiably broken, no harm is done. DH theorists reject the idea that when a violation of this kind is ATC-justified, "morally speaking the case is closed" (De Wijze and Goodwin, 2009: 533). The case, DH theorists argue, is still open. Feelings of "compunction" are more than mere feelings of instrumental value: they point to something real.

Even on the more modest territory of the usefulness of compunction, there are reasons to be sceptical of this Hare-style account (Walzer, 1973: 171-172). How much good can compunction do, as a tool of moral character development, if detached from any meaningful moral truth? The instrumental utility generated by overvaluing rules is dependent on most, if not all, of the actors involved in the discourse believing that the rules possess an innate value that they do not. When actors become aware that there is no innate value to these rules, the benefits from overvaluing are lost. This seems like an especially pertinent criticism in the context of the social functions that rules play.[41]

Of course, threshold deontologists and their allies do not view rules as having only instrumental worth. However, there is still the same practical problem, of an actor being aware that they did nothing *truly* wrong and prematurely moving on from feelings of agent-regret and the like, risking corruption. This is a risk inherent in any attitude that allows the action-guiding content of a deontic constraint to be waived (including DH), but with views such as threshold

---

[40] This point is most apposite in the cases of the severe episodic cases that Walzer's framework is traditionally used to describe. I think it is also accurate when discussing more quotidian betrayals and hypocrisies. Even if it is inappropriate to expect politicians to justify every grimy and unseemly act they do on a diurnal basis, an explanation of why they have chosen to cultivate such habits and dispositions is to be expected. Paul is expected to defend not every single decision he takes while campaigning, for instance, but he should be able to defend *why* he has taken on these attitudes.

[41] Maintaining such rules, then, requires dishonesty and duplicity. This is exactly the kind of dishonesty that DH theorists might point to as being part of the quotidian dirt of politics. The backers of a Hare-style view cannot call on this resource.

deontology there is no moral cost *when we calculate correctly*. This lowers the stakes, weakens the psychological protections, and reduces respect for the deontic constraints.[42] It will be easier to restore these moral rules in normal times if they were not regarded as *completely* suspended even when their action-guiding content was temporarily denied.

Restoring a broken moral rule, in other words, is easier when we recognise the dirt that comes from breaking it. Threshold deontologists thus risk erring closer to a more utilitarian cost-benefit analysis. Acknowledging dirt avoids this and encourages reluctance and introspection, by adding an extra appreciation of the costs of DH beyond the weighing up process itself (De Wijze and Goodwin, 2009: 538). Threshold deontologists will agree with DH theorists that, as Williams (1978: 64) argues, "a habit of reluctance is an essential obstacle against the happy acceptance of the intolerable." However, the acknowledgement of dirt gives DH theorists a far more potent weapon to prevent this happy acceptance than that offered by utilitarians and threshold deontologists.

4.3.1 Rights as Deontic Constraints

The importance of double-counting deontic constraints is especially apparent when these constraints take the form of rights. Rights explain what individuals are entitled to. While some entitlements described as rights may have defeasible or exclusively action-guiding importance, others need freestanding evaluative significance. "Has this right been met" is a separate question from contextually sensitive ones such as "was meeting this right plausible" or "was

---

[42] It also assumes an asymmetry between the options in a dilemma. As discussed earlier, I agree with Nick (2019a) that symmetry is the more attractive view of moral conflicts. A symmetrical view requires there being dirt that attaches both to the ATC-justified outcome and the choice not chosen.

violating this right ATC-justifiable?" Leaving these rights unfulfilled or violated thus creates a moral remainder.

The existence of a moral remainder is not necessarily required for justifying remedial duties towards victims. However, without a true remainder, demands for rectification will likely end up couched in consequentialist or contingent language. Remedy will be justified by and therefore subject to the instrumentalist outcomes. This is the sort of contingency rights discourse is usually meant to prevent. Instead, there should be intrinsic reasons for accountability, apology, or rectification in the aftermath of rights violations. This entails treating rights as having power beyond their immediate action-guiding constraints.

Having one's rights violated is one thing; having one's rights made conditional is quite another. Rights need the character of a "firewall": they are given power by their social "construction" and "achievement" as well as by their normative force (Waldron, 2011: 24). Intrinsic, universal, and inalienable worth seems essential to the construction of natural or human rights and to their claims to underpin human equality. Their social function and normative force depend on them being more than mere social goals that are pursued when ATC-justified. Otherwise, rights become more akin to rules of thumb: useful moral shortcuts but with no freestanding value. Many consequentialists will gladly accept this (for example, Hare, 1972: 175), but for those who otherwise cite deontic constraints as having real moral validity, this conclusion should be resisted.

Rights are phrased in the form of 'a right to X' or 'a right not to endure Y', as opposed to 'a *prima facie* right to X' or 'a right not to endure Y unless there is some overriding societal concern'. The appropriate contingencies tend to be built directly into the right itself, such as a protection against *arbitrary* arrest. When that right is violated, a moral debt is created. There is a moral wrongdoing that should not have occurred and should not, if possible, reoccur. A world without arbitrary arrests is the implied goal of this rights discourse. But what does it

mean to say that we strive for a world without arbitrary arrests *except when justified*? That seems unambitious as a goal, a commitment merely to a utilitarian calculus, one vulnerable to adaptive preferences (Sen, 1985: 14; Nussbaum, 2000: 136-142; Nussbaum, 2011: 54) and special pleading. It is not a rights-based political order. To use the language of the rest of this thesis, it risks presenting rights violations as part of the DC of politics.

Basing a right on some kind of societal need, real or imagined, can thus lead to that need becoming incorporated into the very meaning of that right. The compromise becomes part of the understanding of the right in question. Of course, a critic might retort that a "right against arbitrary detention" *simpliciter* will be vulnerable to adaptive preferences around what constitutes "arbitrary". In the abstract, this is true. In this case, however, the debate is still focused on the right itself—on what is owed to the individual—rather than on what counts as an acceptable cost for society to respect that right.

Another conceptual issue with treating rights as solely action-guiding constraints is that it treats the potential victims of ATC-justified violations as commensurable. Alexander (2000: 904) describes an adapted version of the traditional TBS, wherein we secure the terrorist's surrender by torturing either his innocent brother or his innocent sister. A specific harm befalls the sibling that is chosen for this grim fate, yet neither had a prior right not to be tortured if we believe that that right holds only when ATC-justified. They are treated as moral ballast. The specific moral dirt that comes from violating *their* rights should be acknowledged, to reduce their victimisation.

5. Conclusion

In this chapter, I have defended the doctrine of political dirt against its critics and sceptics. While there are certainly cases where the identification of ATC-justified dirt represents a false

and harmful diagnosis, it does represent a very real set of phenomena within politics. Recognising dirt helps solidify social rules and rights even when their action-guiding content is waived. Recognising dirt gives us a good sense of how dirtying practices, even justified ones, affect both individual actors and the broader polities.

Alternative accounts carry with them their own difficulties. A purist denial of the possibility of ATC-justified dirt cripples the ability of most political agents to pursue even the most morally valuable goals. Specificationism about moral norms can be helpful for moral imagination but taken too far can deny the moral worth inherent within key moral norms and projects. Threshold deontology and consequentialism seem more apt for politics on the face of it yet, among other deficiencies, share a failure to appreciate the importance of moral dirt.

CHAPTER III: BREAKING DOWN POLITICAL DIRT

In this chapter, I turn to what I think is a recurring weakness within DH theorising: treating ATC-justified dirt as if it is a monolithic concept, with all instances sharing the same structure, varying only in quantitative degree. Modifications upon this structure are presented as critiques of DH, or as falling outside of the DH framework altogether. A case either falls within the dirty hands framework or it does not.

This error has existed for as long as DH theorising. Walzer's (1973) research question was in large part why politicians are seen as morally worse than the rest of us. His attempt to answer this question was to focus on discrete choices, distinct moments where a *pro tanto* wrong could justifiably be committed. This attempt to provide a monolithic answer to questions of political immorality results in Walzer analysing disparate cases such as the TBS and Ned through the same DH framework. "To dirty one's hands" is thus treated as a sufficient description for all manner of cases, and for this dirtying to be justifiable necessitates it being the option that represents the lesser evil.

Many valid criticisms have been raised against this approach. These include its limited explanatory scope, limitations to the theory's action-guiding potential, and grouping disanalogous cases together. One particularly prevalent theme within these criticisms is that the orthodox DH view handles cases in far too decontextualised a way. This decontextualised analysis neglects many key considerations: the true moral costs, for both the individual and broader society; that certain skills and dispositions need to already be in place in order to perform a dirtying act; whether one should adopt a specificationist attitude towards the conditions and assumptions that render dirtying acts necessary; and that many instances of ATC-justified dirt cannot be understood as discrete or dramatic cases.

Conventional DH analysis thus stands accused of both making too few and far too many concessions to the griminess of politics. This problem arises, I argue, due to the assumption that all instances of DH can be reduced to a single framework. Instead of treating DH as a single phenomenon, we should view it as representing a series of distinct (potentially overlapping) ways in which agents can justifiably get morally tainted. Some good work has already been done breaking down DH based on severity—that is, the gravity of the moral constraint being broken and therefore the stringency of the demand that justifies its transgression. There are both quantitative and qualitative aspects to this distinction. Loosely, we can define incidents of dirt based on whether the acts they demand are Grimy, Dirty, or Bloody, in ascending order of severity.

As well as differentiating severities of dirt, we should also differentiate based on what renders the acquisition of this dirt ATC-justified. This is the main contribution of this thesis: a proposal to break down dirt along not only along the lines of severity of the wrongful act that is being committed, but also in terms of the source of its ATC-permissibility. The demarcation between Dirty Episodes, Dirty Careers, and Dirty Rules will, I believe, bring much-needed precision to DH discussions, as well as helping fortify the approach against criticisms. As mentioned, the orthodox DH view usually assumes that a dirtying act becomes ATC-justifiable when we can confidently identify it as a lesser evil. This process is apt for Dirty Episodes, but is unsuitable for Dirty Careers and Dirty Rules, where the more quotidian goods of politics can be what is at stake.

1. Breaking Down the Idea of Dirt

DH problems, broadly speaking, concern the ATC-justifiable acquisition of dirt (with dirt being defined using the three conditions laid out in Chapter 1). However, to analyse specific cases, including in determining what is needed to make them ATC-justifiable and what kinds of moral evaluations they ground, we need to add more meat to these conceptual bones. The concept of dirtying dilemmas should be broken down along two lines: firstly, we can distinguish them based on the resulting dirt, by looking at the severity and gravity of the dirt that results from engaging with a given case. This is not an especially novel way of distinguishing between dirtying dilemmas. Nonetheless, given that its neglect leads to the conflation of disparate cases such as the TBS and Paul's, it is one worth reemphasising. The terminology I will adopt is "bloody", "dirty", and "grimy", from the most starkly dirtying to the least severely dirtying.

Secondly, I propose categorising DH conflicts based on the source of what renders dirtying behaviour ATC-justifiable. The distinction here is between Dirty Episodes (DE), Dirty Rules (DR), and Dirty Careers (DC). This is a spectrum, with potential for fuzzy boundaries and hard cases. Dirty Episodes involve a discrete decision (or series of decisions) that represent a break both from what politics "ideally" should be and what the contingent rules of politics have managed to self-justify. Dirty Rules involve the unsavoury practices legitimised or incentivised by the institutions and structures of a given political context, yet do not reach the level of self-justification. Dirty Careers cover the break from ordinary morality that is required by or necessary for (democratic) politics, and the choice to effectively engage with the political realm in the first place. Dirty Careers acts are self-justified by the nature of politics. Because political goals are justified, so are they. Often, though not always, Dirty Careers will be associated with the more grimy acts (for example, Coady, 1990: 263).

The five examples running throughout this thesis all include some degree of ATC-justified wrongdoing, but the severity of the wrongdoing varies dramatically. To torture is much worse than to take a bribe, for instance. To take a bribe is, in turn, worse than anything that Paul is considering doing. To treat the decision to torture as structurally the same as the decision to accept a bribe is, I argue here, the wrong approach. The wrongdoings that Dana, Ned, and Paul are considering are very different. These acts are distinguished in both qualitative and quantitative ways.

One position is to argue that the less extreme acts fall outside the definition of dirt. This would remove Paul and Ned from the DH discussion altogether. For example, Thalos (2018: 185, n 3) argues for a more bounded definition of dirt. Thalos resists views that "[exaggerate] the scope of dirty hands beyond recognition": acts that are part of ordinary politics do not belong within the language of DH. Thalos (2018: 173) does distinguish between "dirt" and "blood". This is not a distinction based on severity, however. Blood describes moral responsibility, whereas dirt describes the phenomenological experience of playing a causal role in negative consequences.

The distinction between moral responsibility and phenomenological experience is sometimes of interest—Thalos uses the example of someone who authorises the torture having bloody hands, the actual torturer having dirty hands. As a whole, though, this approach neglects the conceptual bedrock of ATC-justified dirt shared by the acts that Thalos considers as representing an exaggeration of the scope of DH theorising. Diurnal acts of wrongdoing are

excluded altogether from Thalos's view, despite dirt being a useful concept for analysing them.[43]

DH cases, from TBS to Paul, can alternatively be seen as all sitting along a continuum of awfulness, "ranging from those involving relatively minor moral infractions to cases where there is the commission of terrible moral crimes" (De Wijze, 2018: 132). The infractions involved can be very different, yet they are treated as falling within the same conceptual framework. De Wijze focuses on the most dramatic cases that involve especially serious conflicts between evils, treating them as the "paradigm examples" that best illustrate the utility of DH reasoning. The core structure of the conflict is the same, he contends, for less dramatic examples.

Such a continuum of awfulness can be broken down into different categories. Coady (1990: 263) delineates dirt in terms of the "non-legal" categories of crimes, misdemeanours, and discomforts:

> Murder, theft or torture (crimes); broken promises, public lies or secret understandings with political enemies (misdemeanours); certain sorts of compromises, 'necessary' disownings, advancements or manipulations (discomforts).

The language of sleaze could also be apposite here. Tillyris (2015: 61) emphasises how dirt covers all of politics, even when it is functioning well. In addition to the eliminative emergencies that Walzer focused on, the sleaze running throughout politics also merits discussion. Sleaze has two meanings here. One is the "'morally dubious' but politically 'irrelevant' cases of sleaze and corruption" that constitute abuse of power and self-enrichment (Tillyris, 2018: 113, following Williams, 1978: 55-56). These self-interested and unjustifiably

---

[43] Owing to the possibility of Many Hands, even within the TBS, this distinction between moral responsibility and phenomenological experience should be viewed as a matter of gradient.

immoral acts are not of *direct* interest to DH theorists (Tillyris, 2013: 66, following Mendus, 2009: 1). They are dirtying, but not justifiably so.

The second meaning, which Tillyris focuses on, is how effective political practice *outside of emergency situations* can require dirtying conduct. Thus, Tillyris suggests a different meaning of sleaze (Tillyris, 2019a: 18), where the word refers to violations of positive, or aspirational, morality. This is what captures the "lesser vices" including "hypocrisy, dissimulation and betrayal". Violations of negative, or preventive, morality, such as violence and gross injustice, are referred to as bloody hands. Though I do not think the lesser vices can necessarily be treated as synonymous with violations of positive morality, this distinction between the grossly unacceptable and the merely suberogatory is an important one that hits at a useful purpose.

McDonald (2000: 194-195) distinguishes between dirty and bloody hands, using the term in a different sense to Thalos. DH are acts such as corruption, deceit, and duplicity. The concept of bloody hands covers far graver sins such as mass murder. It is "possible to make whole the victim of dirty hands, but not so the victims of bloody-handed actions" (ibid). A survivor of genocide is broken, McDonald argues, in a way that a victim of fraud is not. As well as individuals, I would also extend this analysis to institutions and the "moral landscape" (De Wijze, 2002: 215), in the sense that a dirty act warps an institutional background in a far less profound way than a bloody one does.[44] Furthermore, bloody acts, if directed towards one's own followers and group members, almost certainly violate whatever fiduciary duties a leader might have towards their members (McDonald, 2000: 195). Not so for DH.

---

[44] An act that is "bloody" for an individual thus may only count as "dirty" for the broader society.

I think Tillyris, Coady, and McDonald are proposing the right kind of structure, in that we locate these acts along the same spectrum but break them down into categories. A tripartite model, which identifies those cases that are less or more severe than "ordinary" DH, will be adopted here. I shall use the adjectives "bloody" (which roughly correlates with Coady's crimes, with the TBS as paradigmatic), "dirty" (roughly correlating with misdemeanours, with Ned and North Takoma as paradigmatic examples), and "grimy" (roughly correlating with discomforts, with Paul as the typical example). These cases all share the bedrock of (ATC-justified) moral pollution, yet there are qualitative and quantitative differences between them.

"Bloody" is the most severe category, covering egregiously evil acts and injustices. Both the victim and the perpetrator are irrevocably tainted by such an act, as are those who directly authorise or enable the perpetrator (thus why the 'bloody' hands of democratic citizens are especially stark). Relationships of trust, if there is a fiduciary duty of some sort between the victim and perpetrator, are forever shattered and can never be restored. The severity and high burden of proof needed to justify such acts mean that they tend to manifest in DE. If they are within a DR or DC context, the moral imperative to enter such careers must be great.

Two clarifications are in order. First, these acts need not be *literally* bloody.[45] An act like mass electoral fraud, for instance, is destructive enough to the shared moral landscape that a similarly high threshold is required in order for it to be ATC-justified. Secondly, not all acts that contribute to or result in such evils being suffered count as "bloody" hands. There needs to be a reasonable degree of culpability and intentionality that lies on the responsible agent, or some direct causal role in the harms being inflicted. These less direct and less culpable roles in

---

[45] Conversely, not all *literally* bloody acts will fall into this category. Proportionate self-defence, for example, does not count as dirtying under my account.

evil are still dirtying, but in a less dramatic sense. Not every contribution to a bloody outcome must itself be bloody: for example, not every civil servant involved in Dana's decision will share blood, at least not to the same degree as she does.

Turning to the other extreme, what is meant by "grimy"? Here, a tangible moral remainder has been created, grounding legitimate moral criticism. There is not only a loss of something of value, but something positively wrongful about the path chosen. However, the cases are far less severe. We are not dealing with *serious* moral violations. The language of evil is wholly inappropriate for the agent, as is the language of destroying the moral landscape. Their wrongdoings are suberogatory and would be so even without their ATC-justifiability.

The grimy-handed agent should rightly be aware of the genuine sleaze and discomforts they have engaged in, but this grime is not a reason to feel overly distraught or morally broken. The victims, insofar as there are victims, are not broken by the individual's actions or complicity (or, if they are, the perpetrators for this damage lie elsewhere). The agent responsible has lowered themselves to the murky conditions and grim requirements of the world in which they find themselves, perhaps becoming a willing part of a causal process responsible for some injustice or greater moral violation, but they have not noticeably worsened it or become wholly corrupted by this exposure.

Dirt represents the middle of this spectrum and the average case. The moral landscape, the functioning of justice, and the rights and dignity of individuals are being threatened, damaged, or disrupted. With dirt, there is a serious kind of transgression, yet not of the kind that represents a long-term disfigurement or scarring. Dirtying political acts represent a break from the norm, a violation both of what politics ideally should be and what its own contingent rules have self-justified, and the kind of transgression that should not be normalised. An obligation has been broken (Coady, 2008: 65), an ought has gone unfulfilled, and a serious determinate wrong can be placed at the feet of the agent. However, the (relative) moral

harmony of the responsible agent can be repaired. The burden of proof for justifying dirtying acts is high and firmly on the responsible agent, yet not as high as the requirements for bloody acts. If an agent erroneously commits a dirty act, they can still be part of the ordinary moral community: an agent who commits an unjustified bloody act should not.

1.2 Episodes, Rules, or Careers

The second and more novel way in which the concept of dirt should be broken down, I maintain, is by the source that causes the violation to be ATC-justifiable. This affects the character of the violation and the context in which it is understood. This is the central argument of this thesis: that cases of ATC-justified dirt within politics can be viewed as falling into the categories of Dirty Episodes, Dirty Rules, and Dirty Careers. This tripartite distinction both protects DH theory against common criticisms and better explains moral conflicts in politics.

Dirty Episodes (DE) represent those cases where a discrete ATC-justified wrongful decision must be made. These decisions need not necessarily be one-offs or static, but they need to be somewhat detachable from this agent's broader moral narrative. DE can be assessed at least in large part on their own, somewhat separable from the overall moral trajectory of the agent's life and the assessment of the institutional framework they operate within. They represent, in other words, a discrete decision where the best course of action (and, in symmetrical cases, all courses of action) inflict moral dirt.

The traditional DH thought experiments, such as the TBS, usually take these forms. Bad moral luck, such as the evil projects of other agents, creates a situation of unavoidable genuine moral conflict. They represent an external imposition, something extraneous disrupting what might otherwise have been, if not morally harmonious, a more tolerable baseline free from this specific kind of predicament.

There is an important choice to be made about the kind of moral life that an agent chooses to lead. The Dirty Careers (DC) phenomenon represents a given role morality (in this case, politics) and the resulting costs of doing business. It represents the inevitable "heat of the kitchen" that those in the profession need to be able to bear. Political success entails adopting morally compromising dispositions, redefining one's relationship to moral virtues, and accepting the self-justified rules of politics, causing moral dirt beyond the level of mere opportunity cost of a life not chosen. Quotidian and diurnal acts of ATC-justified dirt, including those that do not constitute more concrete decisions, are captured by the DC claim. This political role morality cannot feasibly or desirably be changed (or less demandingly, there is no dirt in leaving it intact).

The final category I propose is Dirty Rules (DR). This concerns cases where political activity requires engaging with or perpetuating immoral and corrupting structures. Engaging in permissible or morally important projects requires engaging with or exploiting these corrupting institutions. The terms of the competition (both codified and uncodified) require moral transgressions beyond those that can be described as DC. This is dirtying for all agents playing by these rules. Crucially, the source of a DR conflict is a contingent and remediable structure. These structures can, reasonably, be changed. Indeed, there is a moral duty to fix these systems and a moral wrong in leaving them as is. In the interim, engaging with the structure creates a distinct and discretionary wrongness that cannot be captured by the overall dirt of politics. A non-defeasible negative duty not to use these rules is justifiably broken.

DE cases fall outside the self-justified rules of the agent's vocation. They lie beyond the limits of what agents (assuming a decent baseline of moral luck) expect to face as part of that sphere and whatever constitute the usual standards of success.[46] In contrast, DC and DR

---

[46] Of course, for many political professions, being in a position where one expects to be routinely confronted with DE is part and parcel of the career and its expectations. Here DE and DC can blur, as will be discussed later. Still, each DE act generates discrete dirt.

cases arise from the vocation's self-justified rules. For DR cases, that self-justification is contingent and should be changed. The costs of the institutional arrangement are ultimately undesirable even if the costs of a specific agent's interactions with these rules is justifiable.

A rough-and-ready test for determining whether DR or DC conflicts arise in the context of institutional flaws could be as follows: firstly, do some of the moral stains from participating in an immoral political structure fall on the agent, as opposed to lying at the feet of a culpable antagonist or institution? If they do, an agent is in a DR *and* DC predicament. Secondly, are these moral costs acceptable? Is there a moral imperative to change these institutions, and a moral wrong in leaving them as they are? Where this moral imperative is present, the agent is in a DR situation.[47]

1.3 How These Distinctions Interact

The DE/DR/DC distinction is not synonymous with, or necessarily reliant upon, the gravity/severity distinction. However, *when discussing ordinary democratic politics*, the graver scenarios are more likely to manifest episodically and vice versa. Even the most expansive definition of what constitutes the self-justified ruleset of politics, especially of the immutable DC variety, will not capture the most heinous and shocking of acts, for which DE analysis (and, to a lesser extent, possibly DR analysis) comes into play. At the other end of the spectrum, grime is more likely to be associated with the dirt that attaches to careers or life choices. Episodic instances of grime can arise but are unlikely to be of much interest for political

---

[47] As a result, Paul is in a DC position because, while the virtues that he is being asked to cultivate are unseemly, they are acceptable costs of the political process and there is no moral imperative upon Paul to change the structure of politics in such a way as to get rid of the need for such competitive dispositions. It is neither plausible nor desirable as a political project. The costs Paul is taking on are necessary and, to secure the necessary goods of politics that can only be justly obtained via democratic competition, ATC-justified. In contrast, Edna knows that the virtues she must cultivate are not part of a desirable political system and it would be wrong to leave the system as it is. It is worth clarifying now that this distinction is a scalar rather than binary one.

theorists, as it will either concern relatively minor cases or actors who are not part of the political terrain.

DC and DR can also concern bloodier affairs, but this is rare in democracies. A role morality that includes murdering your opponents may make sense for an environment of civil war, but not so for a well-functioning and decently just democratic system. On this point, I have some sympathy for the arguments put forward by critics of the prospects of democratic DH (or DDH). Ramsay and Shugarman both argue that democratic politics creates a new political order which precludes DH. One part of this rejection of DDH is to say that certain acts are no longer ATC-justifiable by democracies. The claim that, within democracies, the "justifications for lies and violence as rational and realistic responses to political problems are less compelling" (Ramsay, 2000: 40), faces limitations. Democratisation might plausibly reduce the quantity of DE, but not eliminate them altogether. Democratic politics can reduce, but not wholly prevent, instances in which value conflicts become dirtying (Nick, 2019b: 87-91). The self-justified rulesets of a democratic political order will impose strong limitations on certain immoral acts, but as will be discussed in Chapter V, democracies can create their own pressures for dirtying behaviour. Still, the DC and even DR of a democratic politics will be much less likely to be bloody.

A related claim is that, once certain acts become ATC-justified, we leave democratic norms and "turn to war" (Shugarman, 2000: 236). This point is worth considering: if certain acts become normalised as DR or DC, then we are moving beyond the realm of democratic politics (though here I am talking about acts such as serious political violence and oppression. Not even grievous democratic wrongs such as gerrymandering quite reach this level). Similarly, some severe and exceptional DE cases, along the lines of Walzer's supreme emergencies, might require abandoning democracy. Where this line is drawn is not the concern of this thesis, but that possibility is important to note. The self-justified ruleset of politics may

lower the threshold of ATC-permissibility, but there is still only so far that can take us before we exit the category of a functioning democracy.

2. Why These Distinctions are Needed

There are three arguments for why I think it is necessary to break down the idea of ATC-justified dirt. Firstly, there is the argument from decontextualisation: treating dirt monolithically neglects important contextual considerations. This limits scope and prevents a true consideration of a decision's ATC-justifiability. Secondly, a monolithic treatment of the concept of ATC-justified dirt binds dissimilar cases together, imposing the same framework onto disanalogous cases. Not all cases that share the same bedrock of ATC-justifiable dirt can be correctly viewed using the same analytical structure. Thirdly, treating all instances of ATC-justified dirt as conceptually the same at some core level means that there is an inappropriate and undesirable lack of moral imagination in how agents approach these predicaments. This third argument is the practical result of the first and second.

2.1 Decontextualisation

Too many DH theorists, the argument goes, mistakenly conceptualise DH dilemmas as always referring to single, discrete episodic decisions that can be considered in isolation, detached from their broader context. There are two sides to this criticism. The first concerns the nature of the dilemma. Decontextualised theories present DH dilemmas as momentous departures from ordinary politics and moral harmony: they are exclusively presented as being a "rare paradox of action", creating "a single, stark and shaking conflict" that stands out as a "mere

anomaly" in an agent's political life (Tillyris, 2015: 64). The second side of the criticism is that DH theories neglect the world in which the dilemma arose, either by ignoring the background conditions that generated such a dilemma or treating as immutable the terms upon which it arises (Coady, 2008: 89; McDonald, 2000: 188-189).

Coady and Tillyris both advance different criticisms regarding context, though both invite us to consider the broader context.[48] Tillyris focuses on rejecting the idea that political dirt can best be viewed as emanating from discrete aberrational decisions. Coady, meanwhile, emphasises the importance of viewing political decisions in light of their background conditions and broader moral assessments. This can be linked to arguments that bad means can pollute otherwise admirable ends (Shugarman, 2000: 238), that certain moral violations only seem necessary because of the chosen goals (Walzer, 1977: 265-267), or that political agents are prone to conflating their own interests with those of the broader polity (McDonald, 2000: 188).

2.1.1 Why Decontextualisation is a Problem

Decontextualised accounts reduce the purview of DH to the "politics of great gesture", where every moment of dirty hands serves as "a peculiarly shaking, personal and spectacular crisis" (Shklar, 1984: 243). These accounts take as given the prior assumptions, motivations, and other conditions that create a DH conflict, and then treat the DH conflict as if it can be discussed separately from how those factors arose and will continue to persist. Decontextualisation leads to inaccurate portrayals of the moral landscape both before and after a dirtying act has occurred.

---

[48] These criticisms could be treated separately and might even sit in tension. Tillyris focuses on the intractable and perpetual aspects of political conflict (Tillyris, 2015: 64), fuelled by the self-defining nature of internal political goods and the self-referentially justified standards of political excellence (Tillyris, 2016a: 166-167; 2017: 483). Taken as absolutes, these preclude the analysis that Coady regards as necessary to provide proper context. I of course do not take an absolutist interpretation of them.

Part of the problem with decontextualisation concerns reasons of scope. Decontextualisation means that some very common and morally tricky cases that are otherwise classic examples of DH might be excluded from our analysis or incorrectly understood. These include examples such as extrication from unjust decisions, unsavoury compromises common in democratic decision-making and collaborative action, and actions in an environment of moral isolation (Coady, 1990: 268). Examples such as revolutions, where resultant and valuational moral luck play a significant part in determining whether moral violations become ATC-justified, are similarly excluded (Smith, 2018: 205). Similarly, if we focus on DH moments where we aid or abet the "evil projects" of another agent (De Wijze, 2007: 11-16), then reducing our moral culpability solely to the moment of cooperation seems mistaken. The evil project might well continue once it has been started, even when our support is no longer ATC-justified.

Decontextualisation also presents an inaccurate picture of what politicians are like before they encounter a DH dilemma, excluding other potential sources of moral dirt within politics. Decontextualised DH thinking treats these questions as arising for noble politicians who can maintain their uncompromising moral identities up until the critical decision point. This standard is designed for eliminative emergencies and extreme cases (although, even in those cases, a presumption of complete innocence is likely undesirable). Creating the framework around extreme cases, such as the TBS and supreme emergencies, has meant DH thinking is ill-equipped to answer questions about "hustling and lying" that Walzer (1973: 163-164) originally set out to address as well.[49]

DH dilemmas are therefore presented as confronting only the *ex-ante* innocent. The innocent politician must temporarily shed that innocence in the face of a paradox of action, and

---

[49] Given that these more mundane cases of ATC-justified "hustling and lying" still involve a degree of justified wrongdoing that creates a tangible moral remainder, I reject the notion of separating them completely from DH analysis (as done by, for example, thinkers such as Thalos, 2018).

then shall try to reclaim that innocence later once "determinate penalties" have been paid (Walzer, 1973: 178). In reality, far from innocence being the default state for a good politician before they are confronted with a DH dilemma, innocence itself is a choice, something eschewed by effective political actors (Tillyris, 2015: 70-73). Excluding the dilemmas facing already-tainted actors from the scope of DH analysis also unnecessarily reduces the scope of DH analysis by requiring an extraneous baseline level of innocence to exist before DH analysis can be triggered.

Decontextualisation also ignores how DH decisions can be more than a singular act of wrongdoing that warrant a determinate process of repentance. Second-order dilemmas arise in dealing with the repercussions of a dirtying action, such as whether the politician chooses to reveal their dirt (Bellamy, 2010; Tillyris, 2016a: 165-166). Tillyris (2016a: 173) argues that a degree of hypocrisy is needed for the DH politician. *Pace* Walzer, our politicians cannot always own up to their guilt straight away but might need to conceal and prevaricate further to ensure the success of their political goals or avoid negative repercussions of disclosure. They must wear clean gloves to cover up their dirt (Bellamy, 2010). "If [the public] know" their dirt, it is usually a sign that the politician has failed (Hollis, 1982: 389). These politicians are usually then punished by the electorate (Tillyris, 2016b: 15). Decontextualising DH ignores these additional forms of dirt and creates prescriptive accounts that might doom the DH politician's political future by denying political imperatives such as secrecy.

Similarly, a decontextualised view treats politicians as commensurable. It creates a "psychologically impoverished, abstract, and apolitical self" (Tillyris, 2019a: 10), ignoring how we might want different things from different politicians with different political commitments and ideological constraints. Under the traditional presentation of DH, politicians are "static choosers" (Tillyris, 2016a: 166), deciding afresh in each new scenario, whereas both knowing what we should do and knowing how others should assess us requires that broader

story. How we relate to our role will vary from person to person (Mendus, 2009: 118-120), and this affects how we should proceed. Some agents will be best suited to being "moral exemplars", for example (Dovi, 2005: 129).

Other political commitments, including the imperative to remain politically relevant, can impact what an agent should ATC do as well. Sometimes the agent's own desires and long-term career ambitions will seem almost trivial in the weighing up of options, as with supreme emergencies and the TBS. But this is not always the case. Not only can the correct thing to do vary, but what is possible can vary too. Take the TBS. Torture is "not for amateurs" (Shue, 2006: 236), and to do it effectively and proportionately requires skills and the appropriate disposition. Good people and healthy institutions will not have the capabilities and dispositions to torture.

Similarly, a decontextualised view ignores the true *costs* of dirtying one's hands, including in the more eliminative cases such as the TBS. Practices such as torture are prone to "a most dangerous slippery-slope effect" (Meisels, 2008: 157), a "metastatic tendency" to be viewed as "a shortcut" in the future (Shue, 1978: 141-143). Even if we draw clear lines to determine when these abhorrent practices are justified, it can be difficult to "establish decisively the limits [of] relevance" for those criteria (Finlay, 2011: 432-433), and impossible to prevent other actors from using our violations to justify their own (McMahan, 2006: 247; Coady, 2008: 89). We could thus be creating the sort of moral isolation that causes certain DH scenarios to arise in the first place, as well as causing corruption that can be properly understood only in context. Much as we must consider the true costs of a decision by referring to this broader context, so too the ends that we are pursuing need to be contextualised: do they really merit the use of dirtying acts to attain them (Coady, 1990: 265-267)?

Solely because a case falls within the parameters of ATC-justified dirt does not mean it should then be treated alike with all other such cases. A recurring error within the DH literature has been, however, to do exactly that. More specifically, many theorists treat the TBS as a paradigmatic example of DH and then apply the logic of this case, analogously, to other instances (Walzer, 1973, providing an influential instance of this). This approach is flawed. The broader applicability of the TBS should be limited to bloody instances of DE.

What are the features of the TBS that render it inappropriate for broader extrapolation? There is, first and foremost, the gravity of torture. The wrong committed and the evils suffered are so much more severe than the dirt found in the other four examples. This is especially problematic if one believes in a qualitative, as opposed to merely quantitative, distinction between different kinds of wrongful acts. If wrongful acts such as lying and torture exist on the same plane and are separated *only* by the degree of wrongfulness, then it might be possible to use the TBS as a starting analogy. However, if there is a qualitative difference, even a minor one within a mostly quantitative framework, this strategy does not work. Qualitative distinctions break the analogy. The same argument can be made on the other side of the equation, for the rightful ends being pursued. Can different kinds of wrongful outcome (say the bad ends both Dana and the North Takoman gerrymanders are hoping to prevent) truly be treated as commensurable, different only in quantity? Is the logic for the pursuit of good ends the same as that for preventing bad ends? My view is both questions should be answered negatively. Resulting from this, even a minor qualitative difference renders these cases disanalogous.

Secondly, torture is less likely to exist as a practice within a just society. Torture, as discussed previously, requires some institutional preparation: an abusable and corrupting long-

term capacity is created and institutionalised. This is not a wrong that arises in all cases of DH, however. Sometimes a wrongful act does not require this kind of capacity to be developed or maintained, or the capability in question can be part of an acceptable political arrangement.[50] In addition, how we morally analyse the state and its political environment should change based on whether that institutional capacity exists. Any society where a capacity for government-sponsored torture exists is already far removed from a decently just moral baseline. By contrast, acts like lying can exist within decent or morally acceptable societies. Their presence, and the capability to perform them, is not in itself proof of some gross manifest injustice.

In summary, DH theorising has erred in treating the TBS as paradigmatic and as capturing a 'universal' logic that can be extended to all other instances of justified dirt. The TBS involves a great evil and, conversely, requires the presence of a severe moral imperative to override that deontic constraint. This not only prevents the logic of the TBS being applied to more quotidian cases, but the analogy cannot survive being scaled up or down if there is even a slight qualitative difference involved. The gravity of acts such as torture, too, dramatically changes the moral status of the agent who is (or is indeed capable of) authorising it. These dramatic attributes do not apply to many other DH scenarios. This creates qualitative, rather than merely quantitative, differences.[51]

2.3 Action-Guiding Content

---

[50] Most states, even just or decent ones, will likely have the capability to work *with* rights-abusing foreign states or to conceal previous incidents of torture, for instance.

[51] My arguments for qualitative differences will be expounded in the next chapter, where I use discussions of evil as a starting point.

Ultimately, a monolithic depiction of ATC-justified dirt creates weaker and less persuasive normative and evaluative accounts. If someone is insensitive to the perils of decontextualisation, they may leave unexamined the true costs of dirtying one's hands and the supposed necessity to dirty them in the first place. If one treats disanalogous cases as analogous, one will try and apply unsuitable lessons to a dissimilar case. The analysis and prescriptions that might be appropriate for TBS will not all be appropriate for cases such as Ned's. In short, Walzer (1973: 166) was simply wrong when he said that the TBS is simply "a more dramatic example" of the same underlying principles that exist in cases such as Ned's. There are, broadly speaking, costs to treating an instance of DE as DR or DC, and the same goes for the other two categories and the different instances based on severity.

Another key component of this is that the assumptions underlying supposed DH conflicts should always be questioned (McDonald, 2000: 188-189). Firstly, we must consider if the desired end, such as a political candidacy's success or the guarantee of a deal's completion, is actually a good one, and not one whose supposed goodness is instead consciously or unconsciously derived from "a narrower, partisan or even self-serving perspective". Secondly, there might also be cleaner methods of achieving the desired ends, ones that leaders simply are not devoting the necessary "attention, imagination, and energy" to find (ibid). Sufficient contextualisation helps deliver this.

2.3.1 Gravity

Understanding the gravity of a dirtying act provides a framework for appreciating its true costs and, as a result, properly identifying what kind of threshold must be passed before it can become ATC-justified. Mischaracterising grimy acts as bloody will lead to grossly overestimating the moral costs of effective political action; mischaracterising bloody acts as

grimy will lead to a corrosive desensitisation towards the abominable. There are thus dangers to both overstating and understating gravity. The same points can also be made about the outcomes that are being averted. Incorrectly describing a scenario as a supreme emergency would result in immoral behaviours being unjustly committed; not noticing the presence of a supreme emergency when there is one means that options will be unduly limited.

Recall in the previous chapter the arguments for why dirt matters. That an act is dirtying, even if ATC-justified, matters for reasons such as enabling important functions of accountability and social reckoning, understanding impacts on character, preventing happy toleration of the intolerable, and describing moral phenomenology. These are all reasons why the concept of moral dirt can be useful. This utility can be enhanced, and these ends better achieved, by breaking the idea of moral dirt down along the lines of blood, dirt, and grime. A bloody act will have a different moral phenomenology and creates a stronger need for the violated social rule to be reconstructed or reasserted than does a grimy act, even if both acts fall under the umbrella of ATC-justified wrongdoing. Conversely, recognising acts as grimy ensures that agents appreciate both the costs of their actions and the lines not yet crossed.

2.3.2 Source

Understanding the source of a supposed DH act provides a framework for assessing its necessity and how and whether that necessity can be challenged. For example, in DC cases, the reasons for a dirtying act becoming ATC-justifiable cannot feasibly be challenged. This means that it would be inappropriate to use strategies designed for DR, where changing those background conditions is a plausible goal, or for DE, where the source of the conflict is some external immorality over which we might have limited control. Differentiating based on source also helps determine the extent to which the agent could have, or should have, anticipated the

prospect of being faced with the conflict in question, and therefore how prepared they should have been. It also helps agents understand their moral obligations in such contexts. A politician might ask why they should be decent in a world that is so vile (Shklar, 1984: 204). Whether this vileness can be changed or circumvented, and what role the agent's own decency might have in allowing or obstructing said change, will be part of any holistic answer to a question such as this.

Breaking down the monolith of DH thus facilitates moral imagination. Recall the challenge that DH thinking treats as static and fixed the inputs and assumptions that seem to render any given DH scenario as inescapable (Coady, 2008: 89; McDonald, 2000: 188-189). Moral imagination is a necessary part of how we should respond to this challenge (Werhane, 1998; 1999; 2015). Moral imagination means to foster a "critical creativity" about one's own moral development and moral placement (Roca, 2010: 137-139); to try and seek new options when confronted with supposedly simplistic dilemmas via "critical self-reflections". It is to try and see the situation "from different points of view" (Provis, 2005: 293), which are used to inform one's moral reasoning. Moral imagination is helpful for two purposes. Firstly, it is useful for considering what categorisation is appropriate: whether a given dilemma is DE, DR, or DC, in other words. Secondly, and as a result of this, moral imagination is useful for considering what to do when faced with a specific context. Moral imagination allows us to assess whether a dirtying act truly is necessary to achieve a certain outcome and, if it is necessary, whether that necessity is enough to render the act ATC-justified.

Suppose an agent thinks that, while an institutional structure is dirty and corrupt, the participants do not share in the moral responsibility for this. Instead, the moral responsibility is more narrowly allocated, for example falling on the system's designers or a narrowly defined group of rule-setters. The rules are dirty, sure, but they are not in a DR conflict for their own conduct is not dirtying. Moral imagination encourages agents to question this. There might be

a way in which the dirt does fall upon them. Similarly, if an agent suspects that they are in a DC situation, moral imagination invites them to ask if the costs are worth bearing and if morally accessible reform options do exist.

The argument in the previous section, about how the benefits of ATC-justified dirt can be better secured by breaking the idea down further, extends to questions of source as well. For instance, the kind of social reckoning needed for a society that unnecessarily incentivises hypocrisy is different from one that has it as a necessary price for a desirable institutional arrangement. Likewise, the kinds of moral barriers that prevent a too-easy exploitation of and normalisation of DR will be inadequate for DE or DC.

3. Conclusion

Breaking down the concept of dirt helps avoid the pitfalls of decontextualisation, avoids treating disanalogous cases with the same brush, and minimises distortive flaws in evaluation and action-guidance. There are two ways of breaking dirt down that have been identified. The first is in relation to the gravity, or severity, of the dirt that is being attained. ATC-justified moral wrongdoing can represent bloody, dirty, or grimy hands, in descending order of seriousness. The second is in relation to the source, or the characteristics of the supposed conflict and what purportedly creates such a scenario where wrongful behaviour becomes permissible or even obligatory.

The remainder of the thesis will largely be spent elaborating upon the different ways of breaking down political dirt, particularly the distinctions based on the source of the problem. I will do this first by looking at the "traditional" DH view, which roughly correlates to DE. I will then examine DC and, finally, DR. These different views, together, are all useful for explaining real-world political phenomena, including but not limited to cases such as those I set out in the

introduction of this thesis. They should be viewed as complementary parts of the same tripartite

model, rather than as competing explanations of morality within politics.

CHAPTER IV: THE TRADITIONAL DIRTY HANDS VIEW

DH theorising has usually adopted what Tillyris calls an overly "static" view of dirt within politics, presenting instances in which agents commit ATC-justified dirt as dramatic episodes that deviate from regular politics. Tillyris argues that this is wrong, and that owing to its excessively static nature, DH theory misinterprets the true nature of the tension between politics and morality. This is a strong critique, especially in light of the considerations discussed in the previous chapter, though I still think there is a role for this more episodic understanding. In this chapter, I outline the staticity criticism and explore how we should respond to this.

I argue that we should create a delineated space for paradoxes of action, as opposed to paradoxes of character or institutional design, within the DH framework. It is in these instances where episodic portrayals are more appropriate. Not all DH instances should be viewed as aberrational disruptions to what is otherwise an acceptable baseline, or as falling outside of the self-justified rules of politics. Some, however, should be. It is in these instances where a Dirty Episodes framework is most suitable. To support this argument, I also describe in further detail how I believe we should differentiate between grime, dirt, and blood.

1. The "Static" Conception of Dirty Hands

The "orthodox conception" of DH is undesirably "static" (Tillyris, 2015: 63): it views the "conflict between morality and politics as a mere surmountable anomaly", conceptualising it as merely an "incongruence of action" rather than (also) as an "incompatibility of character". I have previously used this force to discuss decontextualization, but its force runs deeper than

that. DH theorists such as Walzer, Tillyris contends, therefore misunderstand the true nature of the relationship between politics and morality. Fundamentally, the charge goes, orthodox DH theorising fails to take the real moral messiness of politics sufficiently seriously. These theorists "focus on a single, stark and shaking conflict that confronts an ex-ante innocent agent", where "unfortunate circumstances" create the anomalous situation of an act being both simultaneously rightful and wrongful (Tillyris, 2015: 64). The true conflict is much more enduring, representing a deep tension between political and moral virtues, and goes beyond this "moralist narrative of temporary disaster" (Tillyris, 2019a: 11).

By treating static cases (such as the TBS) as paradigmatic of DH, theorists commit three broad errors. Firstly, they have fallen prey to "ad hoc sociological [fantasies]" that abstract from the reality of political practice, neglecting how some acts require dispositions to be cultivated (Tillyris, 2015: 65-66). Secondly, these accounts assume that remorse is the only appropriate emotional response to DH predicaments. This imposes too many hindrances for a long-term political career (Tillyris, 2015: 66-67), particularly if the circumstances that led to these choices are mundane or recurring (Tillyris, 2019b). Thirdly, there is also a "paradoxical flavour of optimism" that moral harmony can always be restored (Tillyris, 2015: 67), for both individuals and societies (Tillyris, 2019a: 6).

Instead of focusing on dirt as arising from static choices alone, Tillyris argues for a dynamic account where dirt takes the form of cultivating vices as well as exhibiting them (Tillyris, 2015: 66). This captures "the ubiquitous, diurnal and ordinary dirt" faced by all political actors, including democratic citizens, and does not unnecessarily pathologise necessary conflicts and antipathies (Tillyris, 2019a: 6). "The standards of political excellence" must draw on the resources of politics itself, rather than imposing moralities and abstract concepts that exist outside of politics (Tillyris, 2017: 483). This means conventional morality and political morality are different beasts: "put simply, public virtue and integrity are

incompatible with moral integrity; a virtuous public life is intertwined with the occasional practice of certain moral vices" (Tillyris, 2019a: 4). The orthodox DH view erroneously asks us to focus only on these paradoxes of action rather than the broader question of the tension between different ways of life (Tillyris, 2016a). This is why Tillyris advocates a dynamic view of DH, focusing not on discrete decisions but instead on the costs tied with adopting a political lifestyle, such as adopting private vices as virtues.

Tillyris is not alone in advancing such critiques, though he does provide the most recent and comprehensive synthesis. Williams (1978: 58) and Hollis (1982: 394) both criticised the idea of politicians as static choosers. Mendus (1988: 340; 343) criticised the notion that politicians should be surprised by this kind of dilemma. Johnson (1993: 245) criticised the view that we should focus on explicit choices to the exclusion of dispositions. Hampshire emphasised both the unsuitability of an experienced agent being forever wracked with remorse (Hampshire, 1989: 170) and the intractable and perpetual nature of conflict (Hampshire, 1999: 51). Walzer's DH theorising has been criticised as a doomed attempt to "preserve moral integrity on the slippery slope of power politics" (Sigwart, 2013: 432), or to advocate an implausible "harmony of political interests" (Shklar, 1990: 115). The focus on singular heroic figures means that we end up examining dirt solely through the lens of the discrete episodic choices that individuals face (Shklar, 1984: 243), rather than the implications of the relationships *between* individuals and the unseemly acts needed to sustain basic democratic functions. This focus on "stark choices" represents a caricature of political life (Forrester, 2012: 255).

By breaking down the concept of ATC-justified dirt along the lines of source and severity, we can move away from demanding that DH cases *must* fit within a static episodic structure. As part of this, I defend a space for anomalous paradoxes of action as a (sometimes) useful conceptual framing, without implying, as the orthodox DH view often does, that *all* conflicts political agents face fall into this framework. These arguments follow what I have already said about the structure of choice.

1.1.1 The Structure of Choice

Walzer, like many subsequent DH theorists, viewed DH as resulting from a clash between consequentialism and deontology, namely rifts between "Kantianism and Utilitarian vision[s] of harmony" (Tillyris, 2019a: 10). By presenting DH conflicts as emanating from a tension between competing moral frameworks, these views ignore the broader context. The damage done by violating the rigid and immutable dictates of deontological or private morality is determinate, discrete, and the agent can in the aftermath of the conflict return to honouring both frameworks.

Values are more flexible as units of analysis than entire moral frameworks, and thus less liable to a corrosive staticity. They can be open to change and interpretation, avoiding the charge of immutability that Coady raises, and do not necessarily need to be viewed as "universal" and "sociologically-detached" in the way that Tillyris (2019a: 14) finds objectionable. Despite this, however, Tillyris argues that the VC approach is flawed on this front too. Stocker and De Wijze, he argues, "inherit Walzer's' episodic conception of DH" (ibid), but instead of presenting it as "a stark, momentary, dramatic conflict" between

competing moral frameworks, they instead present competing values in the form of impossible oughts as the source of this momentary conflict. They seem to favour viewing DH in an episodic light, as referring to single choices (Stocker, 1990: 15) and discrete situations (De Wijze, 2005: 456; De Wijze, 2018: 137).

The VC account can be defended against this criticism. Firstly, we can imagine dirtying value conflicts that necessarily entail non-episodic trade-offs. In crafting a space for Democratic DH that explicitly uses the VC model, Nick (2019b: 82-87) highlights cases that involve static and delineable acts of wrongdoing, such as single cases of democratic fraud. However, this account also covers more enduring practices, such as the duplicity surrounding, and the long-term costs of, those single cases. Though Nick does not say as much, this creates space for the second-order DH dilemmas of hypocrisy and duplicity to cover up our discrete wrongdoings that Tillyris (2015: 68) argues is necessary. Conceptually, these second-order dilemmas can be viewed as a value conflict.

De Wijze (2007: 13-14) describes several cases of DH that cannot plausibly be contained as a single decision with discrete effects, such as policies of deception or a long-term abandonment of a value such as justice. Perhaps De Wijze does sometimes treat these decisions in a decontexualised way, but it seems more plausible to treat them as part of some ongoing project, which must figure into the agent's broader political life. This more contextualised interpretation seems compatible with De Wijze's own account of the basic necessary conditions for DH. De Wijze (2018: 132) proposes that we view a necessary condition of DH as stemming from an "unavoidable genuine moral conflict", entailing "justified [moral] violations". There is no reason why the causes or remedies in these situations cannot take the form of chronic or enduring practices.

One aspect of DH theorising that Tillyris (2019a: 8) criticises is "the presumption and language of ex-ante and ex-post innocence and moral goodness". By presenting innocence as the default starting position, even for professional politicians, Walzer's morality assumes harmony is the starting point. Instead, Tillyris favours a more robust value-pluralism that he attributes to Isaiah Berlin, Stuart Hampshire, and Judith Shklar. The "*possibility*, *normality* and *attractiveness* of harmony" are all to be rejected, both at the individual and societal levels (Tillyris, 2019a: 12, his emphasis). Conflict does not necessarily represent dysfunction. Politicians should not strive for moral purity.

Whether a conflict stems from dysfunction requires contextual analysis. Harmony is unfeasible in politics and many important conflicts are unavoidable (thus the DC view), but some episodic instances do represent serious departures from a relevant moral baseline. Adopting a political disposition that allows for considerable distance between political ethics and ordinary morality is neither free of discretionary costs nor constitutes a blank cheque, as we will see in Chapters 5 and 6.

For now, though, it is worth emphasising one potential counterargument. The "dynamic" account of political conflict does not necessarily view the terrain of political conflict as entirely immutable. Politics *can* change, this account acknowledges. However, changing it requires acting separately from the confines of one's ideal politics and accepting the rules as they are. My response to this is to emphasise that acting in such a way, playing within the rules as they are now in order to change them, can still be dirtying in morally and politically relevant ways.

2. What Makes Something Dirtying

Part of the underlying assumption of static conceptions of DH is that the required act is of a particularly jarring and dramatic variety. This is true of some DH acts. Dana and, to a much lesser degree, Ned are being asked to step far outside their moral comfort zones. However, this is not always the case. Grimy acts, episodic or otherwise, do not suitably ground such a harrowing crisis of faith and personal introspection. Quotidian and mundane acts of wrongdoing can add up to, or play a role in, generating serious levels of dirt. However, they do not do this on their own. *Ceteris paribus*, the dirtier the act, the more likely it is to appear as a stark conflict and for it to benefit from the discrete weighing up of costs and benefits that come with DE analysis.

In this section I explore how ATC-justified dirt should be broken down along the lines of gravity, in a way that both captures what is more galling about severe cases and neither denies nor overstates the dirt that comes from more everyday cases of DH. Firstly, I establish the lower threshold for dirt. When does something cross over from ordinary value conflicts into something that merits the language of dirt, even if only grimy? Once this threshold is established, I then ask what separates the ordinary dirty from the merely grimy.

2.1 Moral Costs

For DH to be a distinctive moral category, there must first be some lower boundary of scope that distinguishes a normal moral cost from those that are dirtying. To go back to the three

conditions, when does an unmet moral imperative become so serious as to leave a moral remainder and to ground legitimate moral criticism? There must be some way of working out what kind of acts create important moral residues even when ATC-justified.

Stocker and De Wijze both say that DH analysis kicks in when a certain line has been crossed: dirtying acts are those involving agents being dishonoured or betrayed, their rights being violated, or values that embody these principles being broken. When does wronging an agent tip over into betraying an agent? Is betrayal, in and of itself, the right dividing line? Loosely interpreted it would classify relatively trivial acts of broken trust as dirtying, while leaving out more serious insidious wrongs. De Wijze (2007: 11-16) further attempts to clarify the threshold by defining DH as involving collaboration with or assistance of another agent's "evil project", which narrows the scope but not necessarily in the way that we want. This "external immorality" condition still leaves us with the question of when an immoral project becomes an evil one, or when our support for it becomes sufficient to morally pollute us.[52] My own approach is to focus on breaking the dirt down, as well as drawing its conceptual lower limits, based on gravity.

2.2 Qualitative or Quantitative?

Grimy, dirty, and bloody acts all share the necessary conditions for dirt, but there are qualitative distinction between them.[53] I will use discussions of evil as an analogy here. By showing that evil should be defined qualitatively, I hope these arguments can then be extended to dirt. The

---

[52] Besides, this "external immorality" condition, while often a characteristic condition of DH predicaments, does not seem to be the appropriate dividing line between DH and other moral conflicts. Cases of "lifeboat ethics" prompted by the deliberate actions of other agents, for example, do not become DH if they stem from a deliberate external actor.

[53] The alternative, to distinguish these acts based on quantitative grounds (such as a simple proportionality analysis) and to treat more qualitative distinctions as mere rules of thumb, is also open and I do not think someone who disagrees with me on this point would therefore have to disagree with the rest of my thesis.

lesson I wish to take from them, in short, is that it is a mistake to define evil (or bloody acts, to bring it back to my terminology) as merely the extreme end of a quantitative spectrum. These arguments likewise work in reverse. We cannot define grimy acts as simply being the least bad forms of dirt, nor bloody acts as simply being worse than dirty. The bloodier the act, the more likely DE analysis is to be justified, and so it is important we draw these distinctions accurately.

Definitions of evil are often given in either qualitative or quantitative terms. Quantitative definitions say that an act becomes evil, as opposed to merely wrongful, when it passes a certain threshold of harm inflicted or wrongdoing entailed. For example, Russell (2007: 676) argues for a psychologically thin quantitative distinction, according to which evil acts are those that aim to cause, contribute to, or celebrate what Russell calls "extreme harms". An act goes from merely wrongful to dirtying, and then from dirtying to severely dirtying (or evil) once it crosses this threshold and the harms inflicted or celebrated become extreme.

A wholly quantitative distinction necessarily reduces evil to a mere synonym of 'very wrong' and thus fails to justify the continued conceptual distinctiveness of evil (Calder, 2013: 178-179). In defining extreme harms, Russell's account not only falls prey to the problem of thresholds (Goldberg, 2019: 35), but places these thresholds at a central position in his account—as discussed in Chapter 2 thresholds should not be afforded such centrality. If we end up having too broad a scope for what we mean by evil, then we might see fit to justify the bloodier DH practices and the resulting tools of moral extremity in inappropriate cases. False attributions of evil may be used to justify further wrongdoing (Calder, 2013: 178), thus the danger of using the term 'evil' when applied imprecisely. On the other hand, if we have too narrow a definition of evil, we might become blind to more insidious, normalised, or mundane evils. Additionally, what makes a harm "intolerable, extreme, significant, or life-wrecking" will likely vary (Goldberg, 2019: 39), and be hard to define in solely quantitative terms as a

result. These thresholds are subjective (Goldberg, 2019: 40) and vulnerable to adaptive preferences.

To some degree, these criticisms can be made of all accounts of evil, quantitative, qualitative, or otherwise. However, they are very important considerations here, as defining a certain act as evil has major implications for determining whether the logic of DH applies, and, if it does, then what kind of acts are justified. This is true both with the methods used (what actions should we view as evil, justified only in cases of DH?) and with the circumstances that justify such methods (what choices are so appalling that we must invoke the logic of pursuing the lesser evil?). I think no account of evil, or dirt, can theoretically guarantee an effective and clear answer to these questions, especially at the threshold. Theories of dirt and evil must, instead, have the flexibility, and give us the conceptual tools, to navigate those borderline cases. By making the threshold itself important in defining evil, quantitative accounts give us insufficient tools to step back and ask where the threshold should be.

2.2.1 Qualitative Views

Qualitative accounts instead say that there is a certain quality that evil acts possess that is absent from merely wrongful acts. Following from this, acts attain different degrees of dirt (or a lack thereof) based not on the quantitative intensity of the act, but on distinct qualitative attributes that these acts possess. One common class of qualitative accounts are Perpetrator Accounts (Goldberg, 2019: 36), which locate the necessary conditions of evil in the mentality and motives of the responsible agent. Perpetrator Accounts, however, focus on defining evil *actors*. This is a very useful task in certain contexts (including for questions of punishment and cooperation), but when evaluating acts and outcomes, Perpetrator Accounts are unable to help.

Our focus here is on acts that an otherwise conscientious agent might permissibly or justifiably commit while still being conscious of the reasons against these acts.

Some Perpetrator Accounts struggle less with this limitation, such as Singer's (2004) view of evil as being when one purposely and knowingly does wrong, or Morton's (2004: 57) view of evil acts as those derived from "a strategy or learned procedure" that shows the agent is not constrained by "barriers against considering harming or humiliating others that ought to have been in place". All Perpetrator Accounts, however, face a more fundamental criticism. They fail to capture evil acts performed for banal reasons while instead capturing trivial acts performed with malicious motives (Russell, 2007: 662-664; Goldberg, 2019: 37-38). I find this criticism compelling. Luckily, however, qualitative accounts can take other forms.

De Wijze (2002: 212-213) argues that evil is qualitatively distinct from ordinary wrongdoing. Compared to ordinary wrongs, the dirt is harder to remove, requiring a separate kind of "purification" process, evoking feelings of "horror, disgust and incomprehension" far beyond those that we would feel when faced even with especially harmful acts of ordinary wrongness. We find it incomprehensible how the perpetrators of evil can live with themselves afterwards. To define evil, De Wijze (2002: 215) proposes that we use the Rawlsian notion of reflective equilibrium, striving to make our intuitions about the nature of evil cohere with theoretical principles and vice versa. Qualitative accounts can also include a quantitative component. Liberto and Harrington (2016: 1597-1600) define evil based on a combination of the two, where both the "quality of emphasis" of different factors explains how motive and impact can inform our intuitions of evil. Goldberg (2019: 38) calls for Nuanced Harm accounts, where evil actions require both a culpable perpetrator and a victim as essential components, and both the quality and quantity of the harm committed must be considered.

One especially useful aspect of these accounts is their focus on victims. Card (2002: 16) defines evil as involving intolerable harms that threaten one's ability to live a decent life

generated foreseeably by acts of culpable wrongdoing, Kekes as "undeserved harm" from a culpable agent (1990: 3), and Formosa (2008: 229) as entailing "life-wrecking harm" in conjunction with a culpable human agent. Goldberg argues that a characteristic trait of evil is the exploitation of vulnerability to "give rise to direct dependencies on others to satisfy or to refrain from interfering with the satisfaction of fundamental needs" (Goldberg, 2019: 47); for example, by exploiting a power imbalance to inflict grievous or lethal harm upon another person.

We can thus distinguish between evils suffered and evils committed (Liberto and Harrington, 2016: 1599-1601). Someone can therefore rightfully and coherently claim to have been the victim of an evil, yet the relevant culpable causal agents may not be guilty of an evil. Still, they have played a role in that person suffering evil. Their just cause and moral motivations spare them from the label of an evil actor, but dirt has been generated by the causal culpability for that evil being suffered. The quality of justifiability, countervailing against the intrinsic inexcusability of the wrong that has been suffered, enters the mix. The quality of emphasis is different.

2.3 Defining Grime, Dirt, and Blood

Now we return directly to the task of defining the three categories of ATC-justified wrongdoing as distinguished by severity: grime, dirt, and blood. The previous section discussed how we should define evil. This is useful for the definition of DH in two ways: firstly, in identifying evil actors that it is morally compromising to engage with. Secondly, and more importantly, in identifying (at least some of the) acts that are so intrinsically immoral as to inflict serious dirt even when ATC-justified. I argue that committing evil acts, such as torture and harming

innocents, constitutes the "bloody" extreme of this spectrum, so defining evil, even loosely, indicates how we can go about substantiating that concept.

However, this discussion on evil is useful indirectly as well. The same tools that we should use to define evil can be useful for defining dirt and grime too: a reflective manner to define these terms; attentiveness to victims rather than an insular focus on the perpetrator; and an inclusion of both qualitative and quantitative concerns via quality of emphasis. The distinction between evils suffered and evils committed, too, is an important one. To pursue political power is, at least in part, to accept responsibility for the consequences of one's actions, the actions of one's polity, and the wellbeing of one's constituents. Evil suffered is an important metric here.

The purpose of this section is not to fully define each of the three subcategories of dirt. I instead want to discuss plausible ways in which the distinctions between grime, dirt, and blood might be drawn. These distinctions are important in terms of prioritising discussions of social reform, moral reckoning, and where moral imagination is most needed. They allow the impact on agents and the broader society to be better understood, with where the lines themselves are drawn being a telling object of moral evaluation. They also help flag up the bloodier and more heavily dirtying acts that should, *prima facie*, be stronger candidates for being viewed as a paradox of action.

I will say the least about "dirt". In part, this is because it can be defined negatively in relation to the other two topics. It is not mundane or quotidian in the way that grime is, but nor is it as serious and harrowing as blood. As a concept it captures all the cases of ATC-justified wrongdoing that fall below the threshold of bloody hands yet rise above the level of mere grime. This category covers betrayals, gross lies, serious distortions of just political processes, complicity in some injustices, failure to rectify other injustices, and a host of other serious moral violations. Gerrymandering, bribe-taking, and the like strike me as paradigmatic cases.

What sort of things constitute grimy hands? This category covers more minor violations of morality, especially of positive and aspirational morality. Overall, these acts can be described as suberogatory (Chisholm, 1963; Richards, 1971; Forrester, 1975; and Driver, 1992). Suberogatory acts, the inverse of supererogatory acts, are those that are morally bad yet not prohibited. Grimy acts cover those such as Paul's day-to-day activities. Paul has done nothing monstrous or shocking, but he has lowered himself to the grimy realities of politics in a way that leaves a non-trivial remainder. I am not attached to the terminology "suberogatory".[54] It could be substituted for "minor wrongs", which are still of interest morally speaking (at least cumulatively or when informed by an agent's disposition) yet are not in themselves reason for alarm. I believe griminess needs to go beyond "permissible suboptimality" (McNamara, 2011), however a 'permissible suboptimality that still inflicts dirt' could also ground a definition.

Onto bloody hands. Bloody acts are the most jarring and heinous acts, the ones that are *pro tanto* evil and require the greatest burden of proof to even be considered. Evils committed, even in an environment of moral isolation or perverse terms of competition, are sufficient to create bloody hands. As a working definition of blood, I draw from De Wijze's definition of evil. By understanding the dirt in DH, De Wijze (2002: 211) argues, "we can gain a better insight into the unique conditions of evil actions, when they arise, and the effect they have upon us." Thus we can understand "evil projects" as a necessary condition for De Wijze's view of DH. The structure of my argument runs in the opposite direction, of course, but it is a useful starting point as the connections are already drawn. De Wijze (2002: 218) argues that we can identify three special "conditions or features" about evil, any one of which is sufficient for something to be characteristic of evil:

---

[54] There is a debate about whether the suberogatory represents a true moral category, but that is not necessary to wade into here.

A. There is a deliberate violation of persons with the intention to dehumanise (that is, deny basic respect and dignity to) those powerless to retaliate.

B. The action or project will gratuitously inflict, or bring about, one or more of [what De Wijze terms] "The Great Harms" [by which he means "humiliation, denigration, physical suffering, and even death" (De Wijze, 2002: 217)] to sentient beings with the relevant moral standing.

C. The action or project (or professed morality) seeks to annihilate the "moral landscape."

Under this view, the bloody-handed DH politician is one who assists in bringing about an evil state-of-affairs, conducts an evil action, or assists an evil project but in a manner motivated by moral considerations.[55]

3. The Proper Role of Dirty Episodes

Dirty Episodes (DE) are instances where ATC-justified dirt is inflicted as a result of a (series of) delineable and discrete decision. DE represents the conceptual space where the orthodox notion of DH (that Tillyris rejects) is, if not completely applicable, at least recognisable. A paradox of action, rather than one of character or institutional ideals, creates the conflict, and the source of the conflict can be viewed as an aberrational interruption of what would otherwise be a more tolerable "normal".[56] The dirt from DE cannot be subsumed under the dirt that comes from DR and DC. It is its own thing, standing independently from the self-justified rules of

---

[55] Of course, we do not need to adopt De Wijze's specific view to make space for defining dirt based on DH. There are criticisms of it: Russell argues that De Wijze tries and fails to capture "the paradigms of both malicious evil and banal bureaucratic evil" in his "disjunctive" account (Russell, 2007: 669). There may be merits to these criticisms, but we have a working definition of evil able to define bloody hands. That is good enough for now.

[56] In other words, Tillyris is right to say that DH should not be understood as primarily or only a paradox of action. However, there are instances where emphasising that aspect is the most appropriate prioritisation.

politics. DE are particularly common within politics (especially the higher levels) yet are not necessarily uniquely political in character.

The necessary and jointly sufficient conditions of DE are:

1. Dirt (as defined previously) is inflicted when choosing a justified or rightful action.

2. The moment that inflicts dirt can be understood as a discrete or delineable decision, or a series of such decisions.

3. A determinate wrong is committed, for which determinate guilt can be derived.

4. The dirt in that context can be understood, at least in part, without understanding the agent's political and moral profile.

5. The dirt represents something aberrational, disrupting the previous "default" moral baseline (even if that baseline is far off harmony) and falling outside of what can be self-justified by politics.

3.1 The Space for Paradoxes of Action

Even if we accept the core of the staticity critique, there will still be dilemmas that are best understood as episodic paradoxes of action. Tillyris does not outright deny this possibility. He urges us to reject the notion that ATC-justified dirt is accrued, and innocence strained, "*only* when one is confronted with a paradox of action*" (Tillyris, 2015: 70, his emphasis). There is no reason to assume that this exclusivity applies in the opposite direction.

Creating a role for DE does require abandoning some of the features of the traditional DH model, namely the requirement to restore moral harmony via tools such as disclosure and honesty, the assumption of *ex ante* moral innocence, the lack of sensitivity towards context, any exclusive focus on individual agents, and any pretence of covering all the dirt and grime of politics. These decisions still need to be understood as feeding into a narrative conception

of the self (Tillyris, 2016a: 166, drawing on MacIntyre, 2007): an episodic structure is sometimes acceptable, I propose, provided it can be serialised and understood as part of this narrative. Depending on where it falls in the "narrative", an episodic decision may need to be repeated or concealed, and it may also forever shape the rest of the agent's moral trajectory.

The most important feature distinguishing DE from other instances of dirt is that they represent an identifiable dysfunction. That is, they are not simply an ordinary part of diurnal politics but instead merit *sui generis* treatment. They fall outside the rules of what politics does and can self-justify. One way in which the dynamic account can make room for DE is by using the language of cruelty. Politics often involves lesser wrongdoings such as concealing one's antipathies under a guise of hypocrisy, Tillyris (2016b: 14) argues, drawing on Shklar's (1984) *Ordinary Vices*. Cruelty is rarer. Even if it is required (Tillyris, 2016b: 11), there is a need for an "economy of violence" (Wolin, 2004) with cruelty that is not so pressing with other vices. Hypocrisy is presented as an occupational necessity of politics, practised at least in part to prevent the separate immorality of cruelty from being needed quite so readily. When the need for cruelty does arise, it can be viewed as a dysfunction that requires separate treatment. Hypocrisy, by contrast, is part of the profession.

Another way in which space for paradoxes of action could arise is by using the distinction that McDonald (2000: 190-191) proposes between DH and tied hands. McDonald says that individuals, particularly political leaders, are often asked to tie their own hands to secure grounds for cooperation and prevent the temptations of DH reasoning. These need to be real knots that effectively and meaningfully bind action. One way in which DE arise relates to the question of whether these knots can ever be justifiably broken. In contrast DR and DC arise *within*, or even because of, the scope of tied hands.

Coady (1990: 268) provides three useful examples of cases that can be viewed as DE yet still require appropriate contextual analysis to be properly understood. All are rooted in the

observation that politics involves collaborative action. Firstly, compromises (Coady, 1990: 267-270), where basic principles that serve to constitute an agent's identity are infringed upon. Of course, some degree of compromise is a diurnal professional necessity, but there remain compromises that will be more serious or galling. Waldron (2018) provides a helpful catalogue of a particularly relevant kind of compromise, where the decisions of both one's party and opposing parties can place political actors in paradoxical predicaments. Sometimes sacrosanct political principles and projects stand in tension with one's commitment to the rule of law, for instance. While this is an occupational hazard, the specific decisions merit episodic treatment.

Next, extrication (Coady, 1990: 270-274). Oftentimes, an agent's prior political decisions, or the decisions of their political collaborators, have created an ongoing immoral project or state of affairs, such as an illegal and unjust war. Continued participation in this immoral scheme is morally dirtying. However, there can also be serious moral costs involved in a clean break. Suppose withdrawing from an unjustly occupied country will, predictably, result in even greater evils. It is helpful to imagine these cases as representing an ongoing series of DH decisions, ones that do represent a regrettable deviation from the ordinary costs of politics. This helps to ensure that they are not normalised and that the moral dirt reflected in each option is given its due weight.

Finally, moral isolation (Coady, 1990: 274-279): how to do good when others, unscrupulous or maybe even evil, are setting the rules and standards of politics? Most instances of moral isolation can be viewed as instances of DR or DC rather than DE. However, they can still count as DE under the right conditions, such as when the moral isolation generated is aberrational compared to normal expectations or when the conventions that define DR and DC are themselves in a state of flux. Gerrymandering, prior to institutional normalisation, is an instance of this.

While many instances of ATC-justified dirt defy both the episodic structure and grave moral shock that define the traditional DH view, the more orthodox model still has a role. Some cases that involve an episodic decision cannot be understood as exclusively DE, however. Some acts such as torture *need* to be regarded as ongoing practices as opposed to solely episodic one-offs (Tillyris, 2015: 65-66), even if each specific instance of torture requires a discrete and delineated justification. The choice is not only *should we torture in this case*, but also *are we okay with becoming a torturer?* The stakes of this decision are different and cannot be understood through the static lens alone. There might be an episodic *decision* to set up a *sui generis* torturing capacity, but then the existence of a torturing class becomes a lingering source of corruption within the system, embedded within the narrative of all future decisions.

As the DE analysis suggests, practices such as torture do fall outside of what politics can self-justify. However, they are impossible to perform, at least with the necessary efficacy, as a one-off. There is an implicit assumption—that someone who is trustworthy, in-house, and competent can be called upon to perform the torture—often in the background of TBS discussions, uncommented upon and taken as a given (Brecher, 2007: 21-23). This assumption is not a morally neutral one. To use torture requires the state to cultivate or maintain the capacity to torture or to reach out to an immoral third party. As a practice, torture usually depends upon institutional support. It requires the creation of a "torturing class", as it were (Coady, 2008: 90).

DE cases do not exist in a vacuum and the more shocking and bloody instances can be much harder to justify when examined holistically. This is in part a result of the limitations of scenarios such as the TBS. The choices in these cases are presented to instinctively draw out certain emotional responses from most readers (Coady, 2008: 88; Finlay, 2011: 423-424) and

shut off inevitable empirical complications (Bellamy, 2004: 839; Scarry, 2004: 284). In the end, it is very unlikely that a good polity can preserve this institutional capacity to torture without incurring unacceptable moral costs and using it *only* when ATC-justified.

Instead of the traditional TBS, it might be worth instead looking at cases where we might *cooperate* with a torturer or allow torture to go on with our implicit permission. Similarly, extrication scenarios likely merit more analysis, such as when we inherit a torture infrastructure and face compelling temptations to use it in the meantime. This is an aspect of Walzer's (1973) TBS formulation that warrants greater focus. These cases, while retaining an episodic character, seem more likely to pose relevant questions for real-world debates. They also avoid Coady's decontextualisation challenge, by avoiding "giving the agent narrow choices restricted by external circumstances over which she has no present control" (Coady, 2018: 5). They also allow us to ask, much more coherently, what another agent would do in this situation.

3.2 Moral Emotions

Another implication of the orthodox view is that DH acts should be viewed as a source of suffering. I agree with Tillyris (2019b) that, as a blanket statement, saying that all forms of ATC-justified dirt should lead to suffering is implausible and unattractive. Sometimes tragic remorse is appropriate, sometimes it is not. The distinctions I have offered help to demonstrate this and allow us to move on from the limitations of the orthodox DH view's implications when it comes to moral emotions and phenomenology.

The most common phenomenological view in the DH literature is the one Walzer (1973: 166) endorses. The DH politician *feels* the dirt on her hands. She feels that she is right to try to succeed. In the circumstances this requires dirtying her hands. Her substantive scruples

lead to her being reluctant to do so. She fears betraying those who back her, worries about corrupting herself or her goals, and knows that she will be committing a *pro tanto* wrong. Yet she acts. Crucially, she does not then pretend her hands are clean. She owns up to the dirt. Her rationalisations for her actions avoid "ruthlessly" denying the dilemma by acquiescing entirely to a deontological or a consequentialist assessment, or by making an "Orwellian" attempt to redefine the wrongful act to deny its wrongful character (Lukes, 2006: 5). She accepts, rather than runs from, the messiness of morality and her poor moral luck. The goal, ultimately, is both to avoid "[being] too good for politics", while remaining "good enough" that political success retains its moral value (Walzer, 1973: 168). So far, so accurate.

Walzer's politician will want to clean her hands again afterwards. Penance or punishment must follow. The DH politician commits "a determinate wrong", and must pay "a determinate penalty" (Walzer, 1973: 178) to clean their hands. Just as the dirt in DH goes beyond standard moral trade-offs, the feeling of having DH goes beyond knowing that not all obligations were met and not all values were honoured. By taking a "politically required" action, the DE politician incurs a "*severe* moral blemish", something stronger and more severe than the regrets that come from ordinary moral choices and conventional political activity (Meisels, 2008: 154, my emphasis). It is here where the phenomenology of the orthodox view seems less universally applicable.

What is the appropriate moral emotion for the DE politician? To answer this question, we must first explore the debate around moral emotions in the DH literature. The language of regret fails to capture the agent's causal role, while remorse inaccurately implies that the agent wishes upon reflection that they had not taken the action in question. Given that the agent also has reason to feel remorse if they abstain from DE actions, this creates "an infinite circle of afterthought" (Corvino, 2015: 4). A more plausible answer is "agent-regret" (Williams, 1976: 123). This is a feeling of regret towards past actions that an agent took or participated in, even

when those actions were involuntary, accidental, or unavoidable. It is a way to recognise that even unintentional actions, ones for which agents are not culpable, are causally linked to agents. It distinguishes agents from spectators (Williams, 1994: 69-70). It is the appropriate emotion, for example, for a lorry driver that through no fault of his own ends up killing a pedestrian.[57]

Agent-regret is a useful concept in politics. For example, it can be the appropriate moral emotion for an agent who finds herself stuck within a political structure that creates unjust outcomes yet is powerless to change it or even extricate herself from it.[58] Agent-regret will thus be important in DC and DR discussions later. But it still seems inadequate for DE scenarios. Though the DE politician will regret the circumstances, they stand by *their* action as the ATC-best option available to them. Given the chance, the lorry driver would go back and change his contribution to a tragic chain of events. Not so the DE politician.

De Wijze argues that "tragic-remorse" is the appropriate emotion for DH cases and other incidents of "unavoidable moral wrongdoing" (De Wijze, 2005: 453), and this is appropriate for DE. The agent stands by their actions and their motivations. No fundamental reform of character is necessary, even if there is dirt to be cleansed. The agent still, however, feels shame and guilt. They have still, even if under exceptional duress from others, deliberately broken an important moral rule. The agent was aware of the values that were being violated and understood the wrongful consequences of their choice (De Wijze, 2005: 463-465). The DE politician can thus stand by their acts, yet at the same time be acutely aware of the moral costs.

There are critics of this idea that tragic remorse is the appropriate response. Most prominent is the neo-Machiavellian response. The Machiavellian account is one of the ways of viewing the dirt that Walzer (1973: 175-176) rejects. Machiavelli tells the good man entering

---

[57] Wojtowicz (2022) argues, convincingly to my mind, that agent-regret requires regretting only the agent's own role in what transpired, not necessarily the result itself. This best explains cases where no discrete harm can be identified as resulting from an agent's individual actions, but there are still morally significant consequences for their own moral standing and relationships. This is especially useful when considering "many hands" cases.
[58] Agent-regret is particularly apt for describing instances of circumstantial luck around affiliations, for example.

politics that he will have to learn how not to be good. Doing evil deeds can be justified by the political outcomes and excused by the results, yet they remain *morally* evil. Political and moral judgements thus diverge, and a good man entering politics will have to adopt a lifestyle that prioritises the former. The problem with this view, Walzer argues, is that even though Machiavelli acknowledges the enduring reality and validity of the moral sphere, Machiavellian politicians have no inwardness and are subject only to prudential control. The citizenry wants politicians to be willing to dirty their hands, fully utilising the means and methods of politics, but only for the right reasons and for the right causes. A record of their anguish must be provided. The Machiavellian account does not provide this.

Tillyris (2015: 63) defends the Machiavellian view that Walzer rejects, by saying that Walzer proposes inaccurate start and endpoints. On the usual DH view, dirt arises from the "paradox of action" that renders innocent behaviour unfeasible and undesirable. The politician is *ex ante* innocent, yet is rendered dirty after being confronted with this dramatic decision-making moment. Tragic remorse thus emerges, along with all the associated demands for explanation and reckoning. Tillyris rejects this view. Instead, he argues, the dirt arises from "an incompatibility of character" (ibid). Dirt attaches to an agent because they adopt a political lifestyle; a lifestyle where dirty acts "are neither forced nor unexpected", which requires rejecting "innocence-as-a-disposition" (Tillyris, 2016a: 169). This dirt exists before a conventional DH decision is made. *Ex ante* innocence is thus a fiction. Cleanliness is lost at a prior stage, when choosing to embrace political standards, and thus tragic remorse about losing cleanliness in the face of a specific dilemma is an inaccurately placed emotion at best.

As well as the Walzerian starting point being incorrect, so too is the end-point of Walzerian DH. This is another important point made by Tillyris, and another reason to resist distilling all DH acts to the same framework. Requiring politicians to cleanse their consciences after every DH act will breed either desensitisation to the true nature of dirt, as dirtying acts

are ubiquitous in politics, or cause an "exodus" of good people from politics of the sort that DH theorising aims to prevent (Tillyris, 2019b: 97-98). Many proposed punishments and ways of restoring moral harmony, including by stepping away from politics, go against the fundamental goals of politics such as gaining and securing political office (Tillyris, 2015: 68). Instead, Tillyris (2019b: 106-107) argues that the DH politician should have what Hampshire (1989) describes as "guilty knowledge"—that is, an awareness of the implications of what they have done—but not "guilty feelings". In fact, far from being a suffering servant constantly veering between different crises of conscience, politicians will find there are pleasures to be had in the "excitement" of political intrigue and the "delight" of effectively utilising the skills of politics.

The orthodox DH view on the warranted moral emotions—that is, that they are feelings of anguish, moral suffering, some jarring tearing away from a neater baseline—is inappropriate for many cases of ATC-justified moral dirt (particularly the grimier cases and DC behaviours). It is important to acknowledge and be cognizant of the moral stains in these cases, but it is neither desirable nor feasible to be paralysed with guilt. Similarly, it is not possible or necessarily just for the broader society to react with shock and horror in the aftermath of all such acts, nor should they always call for punishment or other stark forms of socially mandated expiation. Many ATC-justified dirtying acts defy the episodic template that enable these determinate reactions. Grimy acts, too, call for different emotions to those usually associated with the orthodox DH framework. With DE acts in particular, however, tragic remorse will often be appropriate.

4. Conclusion

As Tillyris notes, DH scenarios are generally treated as episodic and harrowing departures from moral harmony. I agree that ATC-justified dirt does not require this episodic treatment. The necessary conditions of ATC-justified dirt lie elsewhere. However, I have argued that there is still a role for the orthodox model of DH, restructured and narrowed down as DE. It does capture some morally interesting and politically important situations and choices, particularly those of a certain gravity or that fall outside the self-justified rules of politics (whether by empirical reality or by reaching the limits of what self-justification can reasonably encompass). Assessment of these cases should still avoid the worst excesses of staticity, such as by thoroughly contextualising cases and recognising that some more enduring practices cannot be viewed through this method.

CHAPTER V: POLITICS AS A DIRTY CAREER

A key source of political dirt is the fact that, in important ways, political activity constitutes a Dirty Career (hereafter DC). In this chapter, I set out what it means for something to be a DC and identify the attributes of politics that render it suitable for such designation. To do so, I will be grouping together under this single conceptual banner a varied series of ideas from across the literature. Many different aspects of politics engender different moral treatment compared to that which is appropriate for most 'ordinary' lives, and these differences can both cause ATC-justified dirt and affect how we view other dirtying choices that an agent might be faced with. I show how this idea can be applied, but also explore its limitations. That politics is a DC is a crucial insight when discussing ATC-justified dirt in politics, but it is not the whole story.

What does it mean to say that an agent has embarked upon a Dirty Career? Firstly, there must be a decision to embrace a new lifestyle, so to speak. The agent must consciously adopt at least some of a profession's self-justified norms of reciprocity, metrics of success, and standards of good behaviour. As a result, the virtues and dispositions that an agent must cultivate are different from ordinary moral virtues. These differing virtues can come from a profession's broader role morality and its necessary purposes, or from the internal goods that are needed to succeed within the political domain. There must be something proactively and positively *dirtying* about the adoption of these different virtues, beyond the opportunity costs of the options not chosen, for it to qualify as DC. Dirtying dispositions and habits can be required, as well as dirtying diurnal acts. This dirtying element is something that is either a necessary or ATC-desirable cost of the career, and the career itself is a permissible (or necessary) one. This creates a new (yet not alien) moral baseline from which the agent should be judged.

The political vocation is best treated as a DC—or perhaps more accurately, a bundle of interrelated and overlapping DCs. All other discussions of political ethics throughout this thesis have had this condition in the background. Other professions might, if they meet these conditions, represent DC in their own way. Politics is not alone in having a bespoke role morality and does not have a monopoly on difficult choices; indeed, much of the discussion here has been informed by professional ethics in areas such as law, business, the military, and other professions. Immoral careers, such as that of a freelance assassin, can count as DC. However, for a dirtying vocation to count as a DC dilemma of the sort that is philosophically interesting, it must be permissible to adopt this dirt. Politics is not the only profession that can satisfy these conditions, but it will be my sole focus here.

Notice that this description contains within it the idea of dirt without encompassing the entirety of DH. The dirt of politics as a vocation goes some way towards explaining why politics is seen, even in relatively just and well-functioning democracies, as a somewhat underhanded or grubby career. Pointing to DE and DR is both insufficient to capture the moral messiness that seems to define politics, and inadequate for defining what might constitute a political role morality. Much as DE or DR are inadequate to explain the ineluctable griminess of politics, so too DC is unable to explain the entirety of what makes politics dirtying. In this chapter, I also consider what I call the Different Worlds (DW) view. This is the idea that politics consists of a different moral world altogether, and thus questions that I call DE and DR should be considered entirely through the lens of what politics self-justifies. Ultimately, I reject this approach, as not all questions of political morality can be subsumed under DC.

1. How Politics Becomes a Dirty Career

There are four interrelated stages to the DC framework that I examine here. The first part of the DC framework is that politics represents a different way of life, one that stands at least

some distance away from ordinary morality. The second is that politics self-justifies many of its rules and standards, and these rules must be accepted to enter politics. The third is to emphasise how the act of governing itself entails certain actions that require a kind of break from ordinary morality. Pluralism and diversity both create situations that require a different disposition, including a willingness to compromise. Similarly, a career in politics requires running towards situations of compromising moral luck. The fourth is that political virtue is noticeably distinct from ordinary virtue. To be a good and admirable politician requires dispositions that are not the same as, and indeed can be at odds with, those that are expected of ordinary citizens.

1.0.1. Ordinary Morality?

I contrast these features of DC with "ordinary" morality, by which I mean the ethical and moral standards that should guide most individuals ("non-political" morality would suffice too). Ordinary morality does not make concessions towards the self-justified goods of politics, does not entail grappling with the most morally vexing demands of governance, and does not require making the same compromises on integrity and personal virtue. For the agent operating within ordinary morality, they not only avoid the distinct dirt that comes from accepting a DC, but their relationship to more episodic instances of ATC-justified dirt remains as normal. They can treat these cases as aberrational, as agonising, as a serious paradox of action. They are not wading into the paradoxes of character that redefine their relationship to such instances.

I do not want to specify what the standards of ordinary morality, as opposed to political morality, should be; only that they exist, that they represent an important moral baseline, and that appropriate political virtues differ in important ways from these ordinary virtues. There are two broad ways to understand this idea of ordinary morality, both of which can serve as an

167

effective foil or contrast to the idea of political morality. The first, and stronger, version is to associate it with moral goodness *simpliciter*. The virtues of innocence, monkish goodness, and adherence to a strict purism can all be incarnations of such a view.

The second, weaker yet more interesting, version is to make room for the "ordinary vices" (Shklar, 1984). In all our lives, we will encounter aspects that require us to abandon innocence-as-a-disposition, to engage in ATC-justified wrongdoing, and to give up the virtues of innocence. The demands of a citizen, then, can be different from the abstract demands of being a good person (Shklar, 1984: 33). Hypocrisy is not only a vice for politicians, but one that we all need to engage in to some extent to maintain our social life and the social fabric (Shklar, 1984: 76-78). A stark division between public and private imperatives, then, cannot be drawn. The private vices that are seen as really constituting 'public' virtues are often tied to having responsibility *for* others instead (Shklar, 1984: 242-243). These questions of character arise everywhere.

It is this second view that poses more of a challenge to my argument in this chapter. If this is true of all of us in some way—say in our capacity as citizens in a liberal democratic regime—then DCs lose much of their distinctiveness. I am not too worried about this potential challenge. All that is needed to define DC is to say that there is a political morality that is distinctively different from what most actors face. Ordinary morality might be dirty, yet political morality is dirtier and dirty in different ways. There are, in both degree and in kind, relevant differences that I hope to illustrate in the sections that follow.

1.1 Politics as a Different Way of Life

The first part of the DC claim is the notion that politics represents an entirely new set of values and a new way of life. Agents who wish to succeed within politics accept this new way of life

and the dispositions and habits that it requires. This is the first (and arguably most important) choice that individual actors must take vis-à-vis their relationship with politics: do they embrace the necessary features of the political way of life and adopt the new moral profile that comes with it? To the degree that the agent answers this choice in the affirmative, the agent has accepted a political role morality and (at least partially) steps outside the action-guiding infrastructure of ordinary morality.

What it means to live in integrity with one's values will be different in political life than it is in ordinary life. For Paul *the citizen* to live in accordance with (say) the democratic value of respecting others in his personal life, the imperative is to avoid *personal* involvement and complicity with disrespecting actions and attitudes. For Paul *the politician* to do that, it requires the active promotion of respect for others as a political good. Political integrity may well require Paul to treat an individual in his personal life with disrespect, if that secures victory in a legislative battle of the kind that better promotes societal expression of respect. Agents cannot properly live by both the political and personal notions of integrity, so must choose which one to prioritise. It is possible to view this choice as *the* choice affecting political morality (and possibly the departure from ordinary morality as an almost absolute one), as the key motivating reason for all the anguish and discomfort that orthodox DH theorists such as Walzer try to explain by using the more episodic model of DH.[59]

A common view in the literature is that the source of dirt in the political lifestyle is thus from a choice made between entirely different moral worlds.[60] The problem for political actors is not that these moral worlds collide and that agents are torn between competing pulls upon them (as the consequentialism-versus-deontology or public-versus-private depictions of DH would argue), but *that agents must instead choose which of the moral worlds to subscribe to*

---

[59] As I will explain later in the chapter, I am critical of such a strong application of this insight, however I believe those who reject it completely are guilty of being undesirably "static", as explained in the previous chapter.

[60] I defend the view in this chapter that at least *some* of the dirt from politics is of this kind but argue that there are limitations to how far this observation can go.

in the first place. It is a choice not between different values, but between "entire systems of value" (Berlin, 1980: 74) and "different ways of life" (Hampshire, 1989: 177). "We lose something of value" whatever path we choose to go down (Mendus, 2009: 88). There is more than just an "incongruence of action", but an "incompatibility of character" (Tillyris, 2015: 63). To pursue an effective political life requires abandoning "innocence-as-a-disposition" (Tillyris, 2016a: 169) to a far greater degree than is required by most lifestyles, even before becoming guilty of actual wrongdoing. This adoption of a political mindset is necessary, going beyond mere technical know-how, to understand "the order and the deeper demands of politics" (Philp, 2007: 52).

This is best understood as a neo-Machiavellian view (Tillyris, 2019b: 104-105). Machiavelli called for aspiring politicians to "learn how not to *be* good", rather than "merely 'learn how not to *act* well'" (Tillyris, 2015: 70, Tillyris's emphasis). For Machiavelli, "being morally good and being politically successful are not just distinct but conflicting aims" (Lyons, 2022: 3). As Berlin (1980: 57) argues, it is permissible to live an innocent monk-like existence focused on one's personal integrity, but this innocent existence is incompatible with political virtue. To take political responsibility while complying with these monkish virtues is to invite irrelevance or destruction.

Effectively pursuing political life entails the suppression of "private qualms" (Berlin, 1980: 59) about this task and the abandonment of moral integrity as traditionally understood (Mendus, 2009: 88). The agent, effectively, chooses the virtues and qualities of a life of "experience" as opposed to a life of "innocence" (Hampshire, 1989: 170). This is not to say integrity itself is rendered meaningless, but it is an integrity that must be politically justified and compatible with political goals.[61] Living a life of experience requires more than merely

---

[61] One justification for this more modest political form of integrity can be that personal integrity is societally important because it shows the agent can be expected to stand by their commitments (Mendus, 2009: 26-27). Integrity thus can have necessarily *political* value. However, this is not an especially strong form of integrity when compared to its more intrinsic or moralistic explanations. This political form of integrity will often not, for

guilty behaviour and guilty knowledge. It requires a certain attitude towards this: the decisions must be taken decisively, if not necessarily with enthusiasm, then at least with no mourning for one's lost innocence.

This view is *not* a full-throttled rejection of traditional morality. It is not saying, as consequentialists would, that concerns about one's adherence to deontic constraints and norms of character lack any intrinsic value. Rather, the position is that politics requires consciously entering a separate sphere and abandoning any claim to live up to the full demands of conventional morality. Goodness must be "thrown away" (Walzer, 1973: 176) to make any progress, rather than merely gambled with. There are "two irreconcilable worlds, one of which must be relinquished" to some degree (Tillyris, 2016a: 172). But goodness, as traditional morality understands it, remains intact as a concept. Machiavelli implores good men to get into politics, as the pursuit of power and glory should not be the exclusive domain of bad men. Traditional morality is important for determining the character of those who might enter politics, and so must remain stable, independent of political norms, to allow this.[62]

1.1.1 Politicians as Suffering Servants?

An alternative way of understanding the choice facing aspiring political agents can be found in Weber's (1919 [1946]) work. Whereas the Machiavellian approach draws two separate 'spheres' into which an agent chooses to enter, the Weberian approach posits two separate moral characters: the ethics of responsibility and the ethics of conviction (or ultimate ends).

---

instance, have anything to say against drawing upon the excuse from alternative cause (Thompson, 1980: 910): when occupying a position that imposes costs on our own personal integrity, can we say that we are minimising harm by denying said position to a more deplorable or damaging agent?

[62] There is also some space within politics for good actions. When callousness is not decisive and there is room to act virtuously, a good man in politics should still take that opportunity provided it does not engender long-term weakness (Machiavelli, 1531: 58). Immorality is, under some interpretations of Machiavelli, the "last best" option, rather than just another option we can choose without hesitation (Skinner, 1978).

Both are important, and both are of moral worth and value, but they are inherently incompatible (Kim, 2022: 6.3). Politicians must choose between them. It is possible to do the right thing by one metric yet completely fail by the other: the use of violence can represent a serious failure for those who seek to live in accordance with pacifistic principles but may be par for the course for the politician who wants to minimise violence committed by others.

For Weber, "the genius or demon of politics lives in an inner tension with the god of love" (Weber, 1919 [1946]: 126), and to seek salvation along the "avenue of politics", where violence and force are innate and intrinsic components, is an impossibility. One could also say that there is a distinction between public and private oughts. While politicians do not *exclusively* deal with public oughts, and nor do they monopolise them (Finlay, 2011: 439-440), they do encounter them more frequently.

On this account, then, politicians do not strictly need a separate moral character altogether, just an acknowledgement that they must operate within separate realms. Politicians must accept that the political career requires navigating between these two existences, dealing with these two competing types of ought. Unlike the Machiavellian politician, the Weberian politician will face a continuous choice about which character they wish to embody at any given moment and which form of excellence to pursue. One, however, is much more suited for political activity and political power, often being an outright necessity.

To achieve success in politics, even when motivated by good causes, is to adopt the ethic of responsibility and abandon the ethic of ultimate ends, to the degree that one must sacrifice one's soul. The Weberian politician is thus a "suffering servant" (Walzer, 1973: 177) or a "guilty martyr" (Rasmussen, 1972: 52), a tragic hero who turns decisively away from salvation and love (Griffin, 1989: 39; Kim, 2022: 6:3). The key difference between Machiavelli and Weber for our purposes is thus that the Weberian politician is suffering, whereas the Machiavellian politician is not.

Weber would prefer to define the sort of character that can best overcome this tension, rather than set out how every politician should navigate this tension. He holds out hope for a "politician with a sense of vocation", the *berufspolitiker*, to combine these two ideals via sheer force of will and to realise their true complementary potential (Weber, 1919: 368). How this synthesis looks in practice is hard to nail down, though it will at least involve a "a sober confrontation with the reality of politics" such as the need for coercion and the threats that come with it (Kim, 2022: 6.3). The Weberian politician is a suffering servant, but if they have particular skills and the appropriate strength of character, a synthesis is possible.

Walzer (1973: 177-180) rejects both the Machiavellian and the Weberian accounts. Once he has set aside conventional morality, the Machiavellian politician is accountable only to the "prudential control" provided by the self-justified goods of politics. They can collapse into being an "unrestrained despot" (Shklar, 1984: 209), with no way for outside observers to detect if they have more worthy goals or if their actions are ATC-worthy of support. The Weberian politician has a laudable self-awareness in contrast, but no means of external expression or accountability for their choices. The Weberian politician's actions are judged internally, not by locating them within a broader framework of public reasons for action. The Weberian language of losing one's soul, too, has costs: it is unclear why and how a politician who relinquishes his soul by severely violating the ethic of ultimate ends can then reclaim it (Walzer, 1973: 178). Ultimately, Walzer favours what he calls a "Catholic" account, inspired by Camus (1949 [1958]), whereby a determinate wrong inflicts a determinate penalty that requires a determinate act of penance. Those who argue that DH acts therefore require punishment tend to have this determinate Camus-inspired model as a background assumption.

The underlying assumption behind Walzer's rejection of these views, however, is that the dirt in politics is best reduced to these episodic moments, with determinate violations that entail determinate consequences. As I have argued, Dirty Episodes do indeed have an important

place in the moral map of dirt. The Machiavellian and Weberian accounts do err in focusing on issues of character and personal virtue to the exclusion of specific choices. There are, as I have also argued, some instances of dirt that cannot be collapsed to these episodic abnormalities or dramatic moments of personal choice. The model of determinate guilt meriting determinate reprobation becomes unsuitable, as it is the adoption of new rulesets and new mindsets that is both creating and justifying the departures from ordinary morality. It is the entry into that separate sphere, the favouring of that different character, that inflicts much of this dirt and defines the agent's moral character. Weighing up each decision discretely is impractical and misses the real force of what is dirtying about the political lifestyle. Assigning determinate guilt can be even trickier owing to the increased 'many hands' nature of dirtying institutions and structures.

The Camus account thus appears insufficient for DC. So how should the politician view their participation in a DC? The Machiavellian account suggests a wholesale adoption of the goods and metrics of the political lifestyle. It is debilitating, and a source of immense practical weakness, to feel anguish. The Weberian account, in contrast, suggests that the politician should be perpetually suffering in the face of these choices, knowing that they still owe fealty to whichever character they are not embracing in the moment, forever striving to become the *berufspolitiker* that can make these two ethics compatible. It is a grim picture of professional ethics, a "moral anguish" approach (Griffin, 1989; Skerker, 2020: 255-257). Which of these two accounts should we favour when it comes to describing the dirt that comes from the political way of life?

Though I do not want to say that either account has a monopoly, I believe that ultimately the Machiavellian account is *closer* to the truth in most DC cases. The "DH as suffering" narrative adopted by both Weberian and Walzerian accounts struggles on two counts (Tillyris, 2019b). Firstly, in treating DH as inherently something that flows from dysfunction, it ignores

dirt that can be seen as stemming from acceptable or even desirable parts of politics. Part of what grounds the appearance of dirt is this idea that something desirable has been lost and a desirable harmonious state has been denied (Tillyris, 2019b: 112-116). Conflict is the paradigmatic example here (Hampshire, 1989). If one treats the manifestation of conflict as always representing a failure of politics, or at least as evidence that politics has not yet succeeded at abolishing conflict, then the results of conflict can be seen as dirtying. "Normal function", defined as the absence of conflict, becomes associated with "serenity and even pleasure" (Edyvane, 2011: 412), but if conflict is seen as not necessarily manifesting from something dysfunctional, the necessity of this association disappears. This is an especially apposite line of argument in the context of DC. The different way of life, and its costs, do not represent something aberrant or flawed.

Because of this, we proceed to the second flawed assumption of the "DH as suffering" narrative: that it is the correct response from politicians. The Weberian politician feels dirt on her hands owing to her inability to reconcile the two ethics, while the Walzerian politician feels dirt that is tied directly to the determinate costs of his actions. The Machiavellian politician, instead, has to a degree internalised the different attitudes and dispositions of politics. What is shocking and shaking for others is not so for them. The difficult choices and balancing acts of politics are something that she can approach with tranquillity or perhaps even excitement, a chance to show off her own mastery of the goods of politics (Tillyris, 2019b: 115-116).

This seems to be a more phenomenologically plausible account too. Not all politicians who handle even exceptional DH choices do so with sadness and remorse. Some politicians respond to the challenges of the political lifestyle with "an appreciation of the demands and standards of excellence intrinsic to [the] way of life" (Tillyris, 2019b: 111), while remaining cognizant of the moral costs of their actions and the imperatives of victory. This seems normatively desirable as well. "A social structure in which the vital outputs of professions are

secured through the moral devastation of those in the profession" (Skerker, 2020: 257) is *prima facie* undesirable, and asking public servants to completely corrupt their characters is an intuitively unreasonable request (Wolgast, 1992: 104; Mendus, 2009: 112). We can minimise this moral devastation not merely by minimising the problematic institutional demands, but by refusing to demand an unnecessarily jarring moral understanding of their roles. Many conflicts in political life cause pain only because they are viewed as dysfunctional and evidence of failures (Edyvane, 2011: 415), not because they actually are.[63]

Of course, I do not wish to argue that the "DH as suffering" view should be abandoned completely. In part this is because denying that DH can entail suffering contradicts what I have argued earlier: the goods and normative assumptions of politics need to be examined, our ideas of necessity tested, the connections between means and ends explored. Being aware of these costs can result in a painful recognition of what has been lost. In part, any role morality of the sort that completely precludes suffering does need to be justified in the name of broader morality (as will be argued for later in this chapter). Too abrupt a compartmentalisation between two spheres can also create adverse effects (Skerker, 2020: 257-260) and cuts a bit too strongly against the notion that political professionals should be able to lead full and healthy lives (Skerker, 2020: 263).

Ultimately, the anguish that a politician feels, and their susceptibility to emotions such as tragic remorse (De Wijze, 2005) and democratic agent-regret (Beerbohm, 2012), depends not only on the moral gravity of what their profession asks of them, but from the source of the dilemma as well. To suffer in the face of some dramatic and deeply abnormal choice is, perhaps, desirable. To feel haunted by the quotidian demands of even a well-functioning democracy's political climate is to suggest that one is not wholly suited for politics. Generally, DC is more likely to entail the Machiavellian model and thus it is probably the most useful

---

[63] This is one of the ways in which a degree of specificationism can be helpful.

model to keep in mind for this chapter. DR is more likely to benefit from the Weberian view of the actors as 'suffering servants', as such cases entail a conflict between what politics is and what it *should* be, and that actors are thus faced with a continuing choice about which ethic to prioritise. DE corresponds most closely to Walzer's model. I do not intend these as total or definitive judgements, however. That politics entails its own separate way of life is the main point of this section and, in contrast to some views, I do not think this need necessarily give rise to moral anguish or even notable discomfort.

1.2 Politics as Self-Justifying Its Own Rules

A key part of what makes politics a separate sphere is the claim that the justifications for many political goods and standards come from within politics. They are self-justified, rather than externally imposed and extraneously understood. When an agent embarks upon a political life, they enter a world of "largely" self-justified rules (Tillyris, 2016a: 163). "The appropriate standards of evaluation arise from within politics, not from an abstract moral standpoint" (Galston, 2010: 387–388). Agents can still pursue morally important goods and projects, but they are doing so by accepting the rules and norms of politics as they find them, endorsing these standards as justified by politics rather than by some pre-political moral baseline. If effective political practice, for instance, requires quotidian lies and betrayals, then that is all the justification a politician needs. This gives political goods and goals a "largely" but not entirely self-justified moral framework (Tillyris, 2016a: 163; 2019a: 18-21).

What leads to politics being a self-justifying realm? At times proponents of this view argue that politics' penchant for self-justification begins where morality ends. A theological example can be found in Luther's account of "the realm of secular power", which has "its own norms and justifications" (MacIntyre, 1966: 127) that often sit outside what religion dictates, with the resulting rules and standards of politics and economics being "self-justifying" (ibid: 123-124). Sometimes an account of politics' self-justifying nature focuses on what the ends of politics are, and these ends by necessity inform all moral rules that follow. MacIntyre (1966: 127) describes Machiavelli as "the Luther of secular power". For Machiavelli, MacIntyre argues, "the ends of social and political life are given" by politics itself. These ends include attaining and maintaining power and ensuring a baseline of "political order and general prosperity". That is why it is important for an independent state to maintain its sovereignty and security, for instance. For individual politicians, securing power is the most direct and secure way to advance one's substantive goals.

Even if one rejects the notion that self-justification begins where morality ends, self-justified goods do emerge. These goods are "internal to the practice" of politics (Philp, 2007: 50). These internal goods are needed to reach goals that have freestanding importance. In addition, politics can play a role in legitimising or defining these values. Political activity can have a creative role in determining what is valued and can be the method through which normative goals are made authoritative for others (Philp, 2007: 83). Whether Paul is successful in the eyes of the broader polity, for instance, hinges on how effectively he convinces others to share his goals and deontic constraints. This means that the "actions and commitments" of other political actors turn these political goods into entities with normative force (Philp, 2007: 52). Though these goals are means in the sense that they enable the achievement or pursuit of other

ends, they are ends in the sense that they obtain their own freestanding importance for those ensconced within politics.

Thus, what is right and virtuous for politicians must be decided in reference to these goals, and these goals themselves can be shaped by political conduct. The "internal goods" that politicians strive for, including the nature of power and the definition of political success, are determined by the way political competition is structured (Tillyris, 2016a: 167) and are necessary to achieve the intrinsically valuable goods to which politics gives rise. Politicians must "accept the authority of those standards" (MacIntyre, 2007: 190). They do this because political goals have a normative weight of their own, with orthodox morality enjoying "neither exclusive rights to morals nor supreme moral status in relation to separate and rival ethical ideals and ways of life" (Lyons, 2022: 3).

It is worth clarifying that not all aspects of the political profession's self-justified ruleset are dirtying. To succeed in politics requires bargaining, trading favours, prioritising relationships with some actors over others, and compromising in all manner of ways (Levy, 2007: 52-53). In and of themselves, these acts do not necessarily entail dirtied hands. There is no intrinsic immorality attached to wanting to succeed in politics (Lawrence, 1990: 426-427; Levin, 1996: 11-15), and, as an often-necessary step in achieving crucially important moral goals, such pursuit of power can even be morally praiseworthy (Walzer, 1973: 165).

However, politics does require acting against the traditional constraints of morality, such as honesty and integrity as usually understood, and cultivating dispositions that run counter to what usually constitutes good moral conduct. There generally must be "a degree of ruthlessness in the pursuit of power" (Philp, 2007: 90). The justification politicians have for doing so runs through the structure of politics. This is not an "anything goes" position (Tillyris, 2013: 254), even before we bring in constraints from ordinary morality. That the norms of politics are self-justifying means both that actions cannot go against "the fundamental ends of

political life, such as a modicum of order, stability and civility" (ibid) and that the vices entailed by politics must be deployed wisely and prudently. Cruelty without purpose remains as unjustifiable as ever.

Another important angle is that certain acts can have their wrongdoing derived from the same source that renders them as politically powerful. In a discussion on the ethics of campaigning, Beerbohm (2016) provides an insightful overview as to why acts such as pandering, mudslinging, and bullshitting should earn our ire. He does this by arguing that campaigning creates an important agency relationship between voter and candidate, a relationship that is distorted by acts such as the above. The wrongdoing in such acts, then, needs to be understood through a political lens and by reference to the self-justified goods of the political way of life. Though Beerbohm does not seek to provide ATC-justifications for such acts, if such justification is to be found, it will likely stem from the same political context that renders it so dirtying in the first place.

1.2.2 Role Morality

Self-justification can be connected to broader empirically rooted arguments for role morality. The distinction between politicians and other actors, instead of relying on some innate or intrinsic distinction, rests on pointing to reasons, rooted in ordinary morality, that show why it is morally beneficial to treat politicians as having different responsibilities and rights.[64] For example, Archard (2013: 782) bases his role morality around a division of labour. The political division of labour is the observation that efficiency is best served when important roles, such

---

[64] Certain iterations of this "common morality" view (Donagan, 1977: 180-189, discussed in Griffin, 1995: 248-254) go so far as to dissolve dirt altogether by saying that common morality covers all the different responsibilities that come with different roles and role morality therefore cannot ever cut against common morality. While I agree with this view in contending that role morality must be anchored in common morality, I reject the notion that there are never tensions for the same reasons that I dispute specificationism and believe that ATC-justified dirt is a coherent concept.

as legislating and governing, are divided throughout the community. These roles, by necessity, carry certain moral prerogatives and duties, and thus a moral division of labour follows the political division of labour. This role morality is not without moral costs, however, and dirt thus arises.

A further way to substantiate or develop this role morality is through the notion of reciprocity: agents within the world of politics decide what the acceptable norms of conduct are for treating each other. For example, if politics involves attempting to defeat rival politicians in elections, this leads to a necessarily aggressive attitude towards one another's job security. Failing to reciprocate would not only be to invite political defeat, but also not to fully embrace the logic of electoral competition that others have agreed upon. Though this aggressive treatment can therefore be internally fair, it can still entail dirt, especially if it affects those who have not signed up for these cutthroat terms.

Can this role morality be defended? It can. Luban (1988; discussed in Griffin, 1995: 227-234) sets out what he terms the Fourfold Root of Sufficient Reasoning to assess the justifiability of a purported role morality. The four criteria (as described by Griffin, 1995: 228) for justifying a role morality's deviation from ordinary morality are that the agent:

"(1) justifies the institution by demonstrating its moral goodness;

(2) justifies the role by appealing to the structure of the institution;

(3) justifies the role obligations by showing that they are essential to the role; and

(4) justifies the role act by showing that the obligations require it."

A political role morality can thus be justified if the agent can justify the moral goodness of politics and the necessity of politics occupying a distinct structure with its own obligations.[65] The obligations of a role, such as that of a politician, thus become "defeasible presumptions"

---

[65] This kind of role morality, then, does not entail rejecting either the comprehensiveness or the dominance of morality generally.

(Luban, 1990: 435), assumed as being the default position yet not immune to the calls of ordinary morality. Of course, this way of defending role morality does not cover all questions that might arise under political morality. For example, DR cannot satisfy these criteria, by definition, as the institutions in question do not have an underlying moral goodness justifying their existence. However, this is an advantage, as I do not think a role morality should be so broad and unquestioning as to accept as given the messier requirements created by structural inadequacies. The exceptions to ordinary morality created by a role morality can "exhaust their ethical content" (Philp, 2007: 55), in other words.

A final note on terminology. Williams (1995) argues that "professional morality" is a more apt term than role morality. This is because role morality tends to encourage a focus on how specific decisions may be different based on one's role, but some acts cannot be performed outside of the context of said role (an ethic of cross-examination of witnesses for non-lawyers, to use Williams's example, makes no sense). Instead, it is the difference in character and dispositions that is important. I agree with the emphasis on character and dispositions that this terminological change suggests but stick to 'role morality' for three reasons. Firstly, it is more common within the literature. Secondly, I do not want to exclude a focus on specific role-based decisions from the analysis. Thirdly, I am unsure whether politics should be described as a singular profession.

1.3 The Dirt in Governing

There are some dirtying features about some of the core aspects of politics itself, namely its practitioners' innate desire to have some control or influence over the exercise of power. These features inflict moral costs that are real yet are necessary parts of the act of government. As participating in government and politics is ATC-justified, then so too is this dirt. What are these

features of governing that are dirtying? In this subsection, I focus on two key features. One is the dirt that comes from operating the machinery of government; the other is the dirt that is created by situations of moral luck that politics inevitably encounters.

1.3.1 The Morality of Governing

Firstly, there is dirt that emanates directly from the acts of governing and politics. Politics is "the art of permissible violence" in terms of the means being wielded, and juggling "inescapable [responsibilities] under partial constraint" in terms of the ends being pursued (Hollis, 1982: 397). There are myriad ways in which this can be articulated. A good starting point is Weberian logic. The means of governing are inexorably tied up with violence and subjugation, and politicians know this. States aim for a monopoly over the legitimate use of force, a monopoly that invariably risks domination (Kim, 2022: 6.1). It is no great stretch to argue that the acts that must be committed or threatened to maintain that status inflict uncancelled moral costs, or that the appropriate disposition for such tasks would be dirtying when compared to ordinary dispositions.

Political leaders also find themselves facing the "tragic irony" of life (Osborne, 2022: 530-531), such as demands for hypocrisy and unavoidable clashes of necessities.[66] Other features of politics can be pointed to in a similar light. Politicians seek to represent the collective, including in arenas such as international relations where normative principles can clash with state interests. Far from stumbling into these positions, politicians seek them out wholeheartedly (Garrett, 1994: 161-163). In pursuing political success, politicians subject

---

[66] These difficulties can run in both directions (Osborne, 2022). The outcomes of politics can be inherently morally tricky: social goods, such as technological progress, can be intrinsically both positive and negative in terms of their impacts on other ends such as liberty. Conversely, how to achieve these political goods can also be inherently morally compromising: for example, to be able to identify a 'general will' of the sort that grounds some justifications for democracy, one might have to engage in hypocritical rhetorical tricks alluding to said general will as if it has already been identified.

themselves to the "stress and frequent indignities" of "the greasy pole" in the process, ruthlessly denying others the same ascent (Garrett, 1994: 173).

Pluralism is another reason for the DC nature of politics. The notion that pluralism leads to dirtying situations, where we do right by one value yet wrong by another, is a familiar one (DeLapp, 2009; also see the thesis's earlier discussion on plural and conflicting values in Chapter 1, Section 2.1). The dirtying impacts of pluralism can be seen in the quotidian functions of governing just as much as it can be in stark episodic clashes, if not more so. Conflict between competing values and interests cannot be eradicated (Galston, 2010: 396). The point I make here, however, is not the familiar one of how politics requires "trading off between different values", nor the related point about how these values can therefore not be treated as absolute deontic constraints (Philp, 2012: 633), though these points are true and support my claim here.[67]

Instead, the point I want to emphasise is that politics is, in modern democracies, about ruling over pluralistic *societies*, with many different competing values, including the politician's own values, some of which they try to make authoritative for others (Philp, 2007: 83). As a result, the task of governing is, at least in part, conflict resolution and mediation in these pluralistic contexts, with institutions designed to both concretise abstract commitments and provide acceptable ways to settle disputes (Galston, 2010: 393-394). Politics is about the "ordering of conflict" (Philp, 2007: 68), or the management of various planes and dimensions of conflict that need to be adjudicated somehow. Sticking to one's own "fundamental ethical commitments" can be impossible when also tasked with governing impartially over and on behalf of a diverse and pluralistic society with clashing moral commitments (Mendus, 2009: 36). Ethical politicians must balance their "substantive goals" with "the obligation to sustain healthy democratic norms and procedures" in pluralist societies (Bagg and Tranvik, 2019: 974;

---

[67] Furthermore, navigating pluralism can also be a task for ordinary citizens in their lives (Shklar, 1984: 245-246).

Waldron, 2018 provides useful examples with regards to tensions between substantive goals, party politics, and the rule of law).

For these reasons, governing a diverse body of people can necessitate some dirt that taints one's character, in the sense that personal integrity must be forfeited and dispositions more suited to conflict and loss must be cultivated. These tasks can be seen as inevitably fraught with practical difficulties as well, grounding additional reasons for dirt. These tasks are conducted with an understanding that perfect fairness is an impossibility (Hampshire, 1999: 39), and that practical reason in politics invariably involves the balancing of greater and lesser evils (Hampshire, 1999: 68, 84)—thus governing requires accepting the fact that treating everyone with perfect fairness and mitigating all evils is an impossible task. Those who govern, or seek to govern, are seeking responsibility for such tasks and the inevitable shortcomings that result from even the effective management of such duties. This is not only a problem for "defective institutions" or for those politicians facing "evil men", but additionally results from "plural institutions and pressures" (Hollis, 1982: 396-397) of even the most desirable of kinds. Messages must be tailored for the right audience, ambitions selectively revealed, and the perfect not allowed to be the enemy of the good.

Because of this plurality of values, interests, and constituencies, politics requires compromise—compromises dealing with our most treasured values, principles, interests, and normative commitments. This ability to compromise on such integral moral goals is a key part of the DC of politics. This is especially so in democracies. Democratic politics requires continuous and sustained "quotidian sacrifices" (Allen, 2004: 39), along with frequent compromises that emerge from the dependencies and conflicts endemic in political life (Margalit, 2012: 157), and it may be that the best way to actualise our values in the political realm is, sometimes, to betray them (Luban, 1985). Beyond merely becoming co-principals of a suboptimal arrangement, these compromises can entail committing wrongful acts and

omitting moral duties (Lepora and Goodin, 2013: 23-26), and frequently result in betrayals of trust (Cullity, 2007: 58). However, this does not mean that every compromise of our values is dirtying. We can distinguish between the evaluative and prescriptive functions of our principles: a particular compromise can be seen as falling short of a valuational ideal while still being the right thing to do for an adherent of said value (Hall, 2022: 220). Refusing to compromise can involve becoming complicit in public harms or passing up a chance to improve an unjust status quo (Hall, 2022: 220-221), or letting down those one is advocating for in an even greater sense than the compromise would have done (Hall, 2022: 224-225).

Still, betrayals, incommensurable losses, and unfulfilled obligations can all result in dirtying compromises. They are "often" dirty (Hall, 2022: 226), with clashing moral imperatives both to engage in compromise and to robustly champion one's representees that can rise to the level of being dirtying. However, the act of compromise is not truly what constitutes the DC of politics. Dirtying compromises could in theory be understood episodically, albeit arising for those with certain obligations and duties owing to their position. It is instead the *disposition* to compromise, and the willingness to search for and adopt dirtying compromises, that predominantly constitutes DC. Instead of interpreting their values and commitments as "thick" ethical constraints that allow for no action-guiding leeway (Hall, 2022: 229), effective politicians must be willing to compromise on them from the very beginning (Hall, 2022: 228).

1.3.2 Moral Luck

Moral luck runs throughout almost every facet of political life. The role of moral luck within questions of high politics has already been discussed in Chapter 1. The success of more adventurous tasks such as rebellion (Smith, 2018) or whether an agent encounters situations

such as the TBS will be a matter of moral luck. Beyond these more dramatic cases, moral luck is endemic at all levels of politics (Nagel, 1976: 142). What makes politics special (though not necessarily unique) is that agents are running towards these situations, willingly striving to adopt positions of leadership in which the responsibility for and dirt of this potentially grave decision-making falls on them. This is one way in which moral luck feeds into politics as a DC: not just in terms of the greater likelihood of incurring situations of DE and DR, but accepting this and the implications thereof from the outset. This is one place where DC and DE interact, which will be elaborated upon later.

Moral luck can also concern whether political engagement in and of itself is justified, and thus whether the dirt incurred in *adopting* a DC ultimately becomes worth it. Williams's (1976) example of the artist Paul Gauguin is instructive here. Williams uses a simplified retelling of Gauguin's life-story, one that makes no claims to historical accuracy, to illustrate how justification can be "essentially retrospective" (Williams, 1976: 119). Gauguin has a choice: to stay with his family in France, or to move to Tahiti to live out his dreams of being a painter. If he opts to move to Tahiti, how we judge Gauguin is determined by how successful he proves to be as a painter. If he becomes a great painter owing to the move, his decision to abandon his old life becomes praiseworthy. The ultimate goal justifying Gauguin's project is something that he cannot know in advance, or that is itself a matter of moral luck about whether he knows or not (Williams, 1976: 121).

There are two types of failures that threaten Gauguin: extrinsic and intrinsic. An extrinsic failure would be an accident that stops his career dead in its tracks. An intrinsic failure would be for Gauguin to arrive in Tahiti, have good circumstantial luck, and then discover that he never had it within him to be a good painter. Both render his travels for naught and mean there is nothing to show for the wrongs committed: however, only intrinsic failure means that he was without justification for ever embarking upon the adventure in the first place. The

consequentialist stakes of extrinsic and intrinsic failure are identical, but what it means for the agent is stark: "the Gauguin who suffers extrinsic failure is merely, like the rest of us, a gambler who sometimes loses: the Gauguin who suffers intrinsic failure is a broken man" (Mendus, 1988: 335).

Of course, there is the question, debated by Williams, Nagel, Mendus, and others, as to whether the luck at question is moral or not, or whether it instead represents another value. What I think is clear is that this sort of luck can be morally important. Even if we do not think Gauguin's artistic realisation itself counts as a moral end, his ability to justify the moral costs that he inflicts upon others is morally relevant. Recall the example of Paul the campaigner. His political success, in and of itself, is not of moral importance. Political considerations may not be analogous to moral concerns, but they do have direct moral significance. Whether Paul can ultimately justify his decision to enter politics will not depend on extrinsic luck, but rather on the intrinsic luck of whether his principles and leadership abilities were worth him eschewing normal virtues and relationships. It is also worth considering how cases such as Gauguin, and thus of moral luck more generally, interact with the notion of symmetry. There is nothing dirtying about Gauguin not choosing to take the risk, as there is no moral obligation for him to move to Tahiti. However, if there are competing obligations on both sides, moral luck can determine which one was the best outcome.

Features intrinsic to the project can lie outside the agent (Williams, 1976: 122), and indeed often do so. Williams here uses the example of pursuing a love affair: the intrinsic success of the project lies in both partners. A political equivalent would be if the intrinsic value of Paul's project was found in the voters. Suppose Paul is fighting for a deregulatory agenda, the value of which depends on voters using that newfound scope of choice well. Paul will have had poor moral luck if the voters turn out to be less morally ideal than he bargained on, or that the success of his political project supposed. The same goes for his political allies, such as the

candidates that he seeks to elect. While knowledge of whether an actor "fits the bill" is not completely a matter of moral luck, there is a degree to which it is, for even the most epistemically gifted and rigorous of agents will have a degree of uncertainty (a degree itself subject to some luck) about whether people will do as hoped.

Is politics different from other careers in this way? In a sense, no, as extrinsic and intrinsic moral luck can affect the ATC-rightness of pursuing any career path. There are however reasons to view politics as different than most careers (albeit by no means unique). Despite its overall permissibility, entering political life requires accepting its self-justified rules and the features that make it a distinct way of life. This is, as I argue throughout this chapter, dirtying, rather than only a lamentable loss of something of value. More fundamentally, politics is a way of life that has a unique claim to concern others and openly uses the resources and freedoms of others in achieving its ends (Williams, 1976: 133). Politics also involves a high degree of inter-reliance with other agents, including the problem of many hands (Thompson, 1980) and the prospect of spattered hands (Rubenstein, 2015). Many hands and spattered hands are not unique to politics, but they are inexorable within politics.

Being equipped with the appropriate dispositions for politics, and having previously adopted the DC path, can also be a matter of moral luck. A wholly idealistic and innocent person thrust into a grievously non-utopian world would be incapable of effective change. Indeed, their innocence could be a recipe for disaster for those around them. Shakespeare's depiction of Henry VI, whose innocence and generosity leads to ruin, is often cited as a literary example of morally good dispositions ill-suited for governance (Johnson, 1993: 245; Tillyris, 2015: 71). To a degree, though, Henry VI is the victim of bad moral luck. That his dispositions led to ruin was a result of him being born into political power, but, more specifically, political power within a certain context: less tumultuous and turbulent times would have mitigated the consequences of failing to adopt the DC disposition. Similarly, the moral force of reasons to

succeed (Gardner, 2004: 81) can also weigh into the relative appropriateness of DC. How important it is that we succeed, versus merely trying, can thus be a matter of moral luck as well. If there is a greater than usual moral imperative to succeed, or particularly vicious adversarial forces waiting in the wings, then an agent without a required DC disposition will have even worse moral luck.

1.4 The Unique Nature of Political Virtue

That politics is a different way of life, with many of its rules being self-justified by internal logic, means that the dispositions and practices that are valued within politics are different to those that are valued in other realms of life. Political virtue is its own category, not detached entirely from ordinary virtue or immune to the critiques that can arise from it, but not tethered to it either. The ideal politician will vary from context to context but will differ from the moral exemplar. This is true even in acceptable and decent political societies. This is because many of the acts inherent within governing and politicking require dispositions of the kind that would be unacceptable or undesirable outside of the political context. Adopting some of these virtues *can* be dirtying (and the political disposition is very likely to be so), yet this is an ATC-justified dirt owing to the permissibility of political endeavours.

The 'ideal politician' is a different entity to the exemplar in most moral traditions. Again, this is a neo-Machiavellian insight, a modern formulation of his writings on *virtù*. To succeed in politics, and to maintain the political goods needed for political success, means that one must engage in hypocritical concealment of one's true positions and antipathies, to trade away one's most substantial convictions to secure consequentialist benefit. "Public virtue and integrity are incompatible with moral integrity; a virtuous public life is intertwined with the *occasional* practice of certain moral vices" (Tillyris, 2019a: 4).

While this is something that we must all reckon with to the degree that we are involved in public life (Shklar, 1984), it is all the more true for politicians. Moreover, the traits associated with "brutality and meanness" in private life are those that can make a successful politician (Shklar, 1984: 71). How these vices are treated, too, is different in the political sphere. As is often the case, hypocrisy serves as a particularly illustrative example. Hypocrisy, as Shklar contends, is something that can be seen as a social necessity for the citizenry as much as it is for the political class. Despite its ubiquity, the charge of hypocrisy is a powerful attack within politics. Political debate thus becomes a competitive game of mutually unmasking one another's hypocrisies, and as a result "the general level of sham" in politics rises (Shklar, 1984: 67).

The more one enters political life, the more must develop these different skillsets, and adjust their ambitions around this reality. The political leader must adopt "the qualities of a lion and a fox", force and fraud, and practise "the leadership qualities of resilience, guile, flexibility, decisiveness, courage, and the ability to master the capriciousness of luck or *fortuna*" (Lyons, 2022: 3). They must value the "internal goods" of politics for their own sake and not abandon them in the name of external goods (Philp, 2007: 50). In addition, they must "adopt a different view of the more conventional virtues that have a place in political life such as clemency and generosity" (ibid). The virtuous politician must "inhabit" the system (Philp, 2007: 52), be able to manipulate his character and mimic other virtues (Philp, 2007: 41), and understand the limits that arise on their means and ends through political rather than solely moral lenses (Philp, 2007: 39; 87).

In terms of the virtues they promote in society, another proposed distinction is that politicians have a greater responsibility to promote "artificial virtues", those attributes that achieve their goodness only because they arise as part of an overall scheme of justice (Finlay, 2011: 434-435). An exclusive focus on "natural virtues", those that are good in themselves

without the "mediating" effect of other institutions, is unsuited for politics. For example, politicians must sometimes violate a natural virtue of beneficence if doing so helps maintain a more important social institution. This might entail refusing an act of clemency if it threatens the rule of law or an act of generosity if it threatens the effective functioning of distributive justice, even if it would be the morally right thing for a private citizen to do.

Let us return now to the example of Paul from the introductory chapter. Paul has two broad choices in terms of what dispositions he decides to cultivate. The individual decisions facing Paul are not a helpful way of viewing his choice, because focusing on discrete decisions distracts us from the real question of moral significance regarding the type of person Paul becomes. Besides, an agonising choice-by-choice approach is a recipe for paralysis. The neo-Machiavellian answer is to tell Paul to accept that he must pursue different qualities and a different relationship to the moral virtues. If he chooses to live in accordance with his old habits and dispositions, without accepting the new standards of goodness and metrics for success that politicians strive for, he is inviting disaster and irrelevance. He is becoming a (far less dramatic) democratic equivalent to Henry VI (Tillyris, 2015: 71).

Traditional morality is not rendered meaningless by this choice. There is something to be said in praise of the attributes that render Paul ineffective. Paul might be violating the demands of political morality (whether this is understood as a subset of broader morality or as a competing form of morality), but there is still something admirable in his conduct. Curzer (2006: 34) explains this by arguing that there is a divergence between duty and virtue. In sticking to his moral guns, someone like Paul (or Dana, as Curzer's discussion focuses on the TBS) is committing "admirable immorality". His decision to remain pure is wrong (as would Dana's, Curzer proposes, if she refuses to torture), but it still represents a morally virtuous instinct (in Dana's case, a virtuous opposition to torture; in Paul's case, a virtuous desire to be honest and truthful).

How is pursuing a political set of virtues dirtying? Part of the answer comes from pointing to the evaluations and standards created by ordinary morality, as these still exist for political actors and are not cancelled altogether when the agent in question chooses another path. Divorced from specific instances and practices of dirt, however, why should this be viewed as dirtying, as opposed to just another opportunity cost?

To answer this question, first we must unpack what is meant by "virtue". To have a virtue is to have an "excellent trait of character" in the form of a "well-entrenched" disposition that "goes all the way down" in an agent's mindset and identity (Hursthouse and Pettigrove, 2016: 1.1). Virtues rarely exist perfectly or impeccably within a person, more often being a matter of degree (Athanassoulis, 2000), and should be judged holistically rather than on a single discrete action (Sreenivasan, 2002). Virtues are more than just a reason for action: they represent "the way that I have made myself, the way that I have chosen to be, and in deciding [to act] in accordance with it, I endorse the way that I have become" (Annas, 1993: 51; discussed in Philp, 2007: 86). This creates a meaning of virtue suitable for both "ordinary" and political virtues.

So why are political virtues occasionally dirtying? They represent the incorporation of the messy aspects of politics into one's character, inhabiting both the contingent and the necessary flaws of the political climate. Practices such as hypocrisy (Tillyris, 2016b) and ruthlessness (Philp, 2007: 90), which constitute vices by ordinary standards of decency and integrity, become ingrained in an agent's behaviour. The prioritisation of political goals and relationships can conflict with other "natural" virtues such as friendship and duties to one's

family. There is also the risk of corruption. The risks of corruption and broader impacts on character have been discussed in Chapter 2, and those arguments are especially apposite here.

Political virtues such as ruthlessness and insincerity are unlikely to be contained easily. They will arise beyond discrete occasions where they can, if evaluated on this episodic basis, be judged as permissible. This is because they represent a "settled disposition" (Philp, 2012: 644) to do something as opposed to a mere capability to do it, having a pull on the agent even if it is not action-guiding or overriding of all other considerations (Philp, 2012: 642). The agent is both able to spot opportunities to deploy these unsavoury skills and has the mental wherewithal to utilise them without too much hesitation or paralysis. A disposition to lie only when justified thus has a "metastatic tendency" to become "a shortcut" to solve other problems and achieve other goods in the same way that a disposition to torture only in extreme circumstances has (Shue, 1978: 141-143). A disposition to lie when justified often collapses into a disposition to lie *simpliciter*. The disposition to lie is, of course, much easier to ATC justify than the disposition to torture. That does not render it without costs.

1.5 What Makes Politics Dirtying

In this section, I have argued that politics can, even in a well-functioning state, constitute a Dirty Career. The self-justified rules and goals of competitive politics, the dirt found in governing due (among other considerations) to the inevitability of pluralism and the heightened risk of bad moral luck, and the tensions between political excellence and ordinary moral excellence all create a bespoke political role morality. As a result of these attributes, political virtue is a different entity to ordinary virtue. A virtuous politician will display different traits and attributes to an ordinary virtuous person, with a greater emphasis on compromise and ruthlessness. While some of these virtues might be contextually dependent and the result of

something dysfunctional about their circumstances (as with DR or a recurring series of DE events), these virtues can also be impacted by DC conditions found in ordinary and well-functioning politics. This is *why* politics can become a DC, and what it means for that to happen. But how can we tell if a politician's claim to be operating within the confines of a DC is a valid one, and how can we distinguish this from other phenomena?

2. What Dirty Careers Cover

In this section, I will set out the necessary conditions for a life choice to count as a Dirty Career. These necessary conditions will be for the concept of a DC more broadly, as opposed to a specific focus on the DC of politics. Of course, politics satisfies these conditions. There is an ATC-justified moral remainder that comes from subscribing wholesale to the political way of life, and this does not mean that the political way of life is institutionally broken. To show how this works in practice, I return to the five examples that I set out in the introductory chapter. The DC framework is useful in properly contextualising all these cases, as well as describing the choice faced by Paul.

As to how DC relates to the other forms of dirt, I conceptualise them as distinct yet not necessarily mutually exclusive categories. There can be overlap or ambiguous cases. In this section I explore the relationship between DE and DC. That many politicians will have to be prepared for, or will inevitably face, many instances of DE is arguably an integral part of their DC. The circumstances of the DC that they have adopted also affect the second-order dilemmas flowing from a DE, as well as what options can even be pursued in a DE. The task of differentiating DC from DR and exploring the many ways in which those two concepts interact and overlap will be reserved for the next chapter.

For something to be a Dirty Career, the following necessary and sufficient conditions must be met:

1. Effectively entering a career requires adopting a different set of values and virtues that break notably from ordinary morality. It is a different way of life. New moral baselines are created as a result.

2. These differences arise from the role's broader purpose and justification, and/or because of internal self-justification that arises from the career's internal goods and practices.

3. There is something dirtying about choosing to pursue this career. This can include accepting the quotidian use of certain dirtying acts. A moral remainder is created, and the agent's moral character is changed.

4. This dirt is all-things-considered permissible. Entering a DC is still something that moral agents can (and in some cases, *should*) do.[68]

This is a definition of DC that is not *exclusive* to politics, though it is certainly built with politics in mind as the paradigmatic example.

The results of adopting a DC are threefold: firstly, it affects what counts as a Dirty Episode for an agent. There is a change in the moral "baseline", with some acts that previously merited conscious deliberation and a degree of caution and determinate regret instead becoming quotidian, if not outright virtuous. An example is trying to get someone of integrity and competence fired from their livelihood. An act like this would be quite an abrupt departure from the norm for most people; it is however something that most politicians in functioning democracies need to be capable of doing. Even though these efforts are not dirtying *per se*,

---

[68] Thus my arguments for the moral permissibility of entering politics, made in Chapter 2 in contrast to purists, are relevant here.

they would be dirtying in most areas of life, absent the reasons and role morality of politics, and require some deviation from conventional morality.

Secondly, it changes how we should view the moral character of agents facing other categories of moral dirt. The virtues that they value, the relationships that they cherish, and the relevant definitions of integrity will be different for politicians. How clean an agent is before an instance of DE will be different based on the demands of their DC, as will the question of how clean they should try to make themselves afterwards. There will be some acts, dirtying for most individuals, that the DC agent does on a quotidian basis, and thus seeking to wash that dirt away is an unrealistic demand.

Similarly, the implications of DE and DR acts will be determined at least in part by DC. For these cases, DC will help determine what after-the-fact disclosure is required, what institutional infrastructure is in place to allow such transgressive behaviour, and the like. Some acts that go beyond the norms of DC, and thus require the DR or DE treatment, will have their *aftereffects* (as well as whatever preparation is in place beforehand) determined entirely by DC. Most questions of state security can be viewed as falling within this lens: Dana would need to consider whether divulging what she has authorised will compromise future state security, rather than focusing entirely on her own moral cleansing. It is, of course, possible that a DC's demands cannot cover the surrounding circumstances of some DE acts, and that ATC-justified practices and questions of disclosure surrounding these acts cannot be governed *entirely* by the demands of DC. Under these more limited circumstances, there will still be a role for DC, however. As well as asking what a moral agent must do when confronted with an especially stark dilemma, politicians must ask what a good politician must do too.

Thirdly, there can also be dirt that is directly generated by the actions and dispositions required for a DC. Viewing this dirt on a decision-by-decision basis would be analytically weak owing to its decontextualised nature, as well as missing out on some of the more important

accumulative and personal costs. Consideration of these longer-term costs requires a more long-term framework of the kind provided by DC. These costs go beyond mere opportunity costs, such as the lamentable closing of doors and the forgoing of valuable opportunities of the type that define almost every life-plan. DCs also require stepping away from the ordinary moral baseline of a good citizen. This includes adopting vices as virtues, displaying greater quantities of vices such as hypocrisy and partiality, and normalising otherwise wrongful acts.

One important clarification: to say that politics is a DC does not mean that all politicians enter the same DC. Both the broader context and the demands of the role itself have an impact here. Some politicians may not enter a DC at all, whether out of choice (whether from naivete or from a determination to serve as a moral exemplar and utilise moral capital) or because the specific demands of their role and their context involves nothing dirtying (if Paul finds that he can contribute to political activity via handing out benign and truthful campaign literature that he agrees with, it is unlikely he leaves ordinary morality). The demands for how an agent must deviate from ordinary practice may not result in any noticeable dirt falling upon the agent.

To illustrate this, consider the kinds of people who are active in the political arena: lobbyists, campaign strategists, legislators, executive officeholders, civil servants, parliamentary staff, party researchers, and devoted campaign volunteers all enter the political realm. They enter the political realm differently, however, and for different purposes. The appropriate role morality varies. This can also change based on the level of politics at which they enter—a local councillor and a head of government face very different circumstances and expectations—or the context that their polity is in—a superpower versus a small city-state, a peaceful developed nation versus a poorer war-torn one.

Allow me to return to the five examples laid out in the introductory chapter. I believe they all support the DC framework in some way, even if the core dilemma contained within them is not plausibly captured as part of a DC choice. Even in the instances where these choices are not covered by the DC observation, the presence of DC as a background condition helps us to better understand what is at stake and how to understand these cases.

Firstly, consider Dana and TBS. The decision itself is not covered by DC. Torture is too severe and grave a practice to be eligible for normalisation via role morality: the moral violation involved warrants determinate treatment every time, falling outside the self-justified norms of any ATC-justifiable role morality. For something as abhorrent as torture to be ATC-justified means that something is seriously wrong: acts like torture should not even be conceivable in functioning or well-ordered polities. That it might count as a lesser evil means that something is *very* wrong. However, Dana is knowingly and consciously adopting a position with responsibility for these decisions. She is accepting a high degree of moral luck: indeed, it would be exceptional were she not faced with DE predicaments. She is also accepting the responsibility to be in situations of moral isolation and to inherit the unjust practices and institutions of her predecessors, which includes the institutional capacity to torture. She is, in other words, volunteering to face the challenges of extrication. The specific decisions will best be understood as DE. Still, the dispositions that she must nurture and cultivate are better understood as DC.

This also gives rise to various second-order dilemmas where the impacts of DC are evident. The orthodox DH view, as espoused by Walzer, requires Dana to fully disclose what she has done and seek public redemption. It also requires her to be upfront about her views on torture, including the fact that she was sincere in her opposition in the first place and yet she reached the limits of that opposition. These second-order dilemmas can, depending on their

impact, sometimes take on an episodic flavour. More likely, they will need to be more diurnal and quotidian; less of a deliberate decision to adopt a certain practice, and more a matter of adjusting her settled dispositions and letting these decisions be made by reference to the internal goods of politics.

Next, take Ned and Edna. The choice facing Ned is a DE one, where an important moral good requires committing a tangible moral wrong. Superficially, the choice facing Edna can be viewed in a similar light, but properly contextualising it instead reveals it as a DR one (as will be explained next chapter), since the contingent structural elements of her institutional surroundings give rise to the form of the dilemma. Edna is being faced with the demands of a role morality, but it is a role morality that she knows is in serious need of reform. However, there is a crucial element of DC theorising in understanding these cases: in understanding the political goods that motivate these decisions, in the second-order vices such as hypocrisy and political ruthlessness needed to make these deals worthwhile, and in the steeliness needed to enter this grubby sphere in the first place. The same applies with the North Takoma gerrymander.

The DC framework is most directly applicable for Paul. It is the very substance of his dilemma. The question facing Paul is what kind of person to become. While he does not want to become a mere functionary of party politics, or a conduit for the moral flaws of politics, effective engagement requires adopting the political worldview and dispositions and making (potentially risky and counterproductive) sacrifices of integrity. Paul is aware of the costs of embarking upon a DC, in other words. His choice is whether or not he wants to enter a DC in the first place, and whether he wants to adhere to his values or advance them politically. Paul also wants to be conscious of the limits of what this DC justifies, however. He does not want to be worse than he needs to be or to adopt dispositions and practices that cannot be ATC-

justified as part of an ordinary political career. Many of the tools he will learn how to use will be vulnerable to abuse.[69]

2.3 Distinguishing Dirty Careers from Dirty Episodes

There are interesting cases where DC and other categories of dirt interact, such as where an instance of DE cannot be understood properly without referring to the DC in the background, and vice versa. To understand how these cases arise, it is important to clarify that I understand these different categories of dirt as being capable of overlap. The overlaps between the concepts need not necessarily be large, but they do exist. Emergency preparedness is a clear example of where this kind of blurring arises. The individual decisions should be correctly understood as DE. However, accepting and pursuing a career with responsibility for such decisions being part of the nature of the office will entail the cultivation of certain skills and dispositions, as well as the development of certain institutional practices. The propensity of politicians to find themselves facing a particularly severe kind of DE, in other words, is part of what creates a DC.

Let us take Dana and the TBS as an example. In seeking and winning high political office, Dana knew that she would be taking on a special moral responsibility for her state's security and control over key institutions of political force. She knew that she would be placing herself in the situation where, if the TBS choice did arise, dirt from the decision would fall on her. Dana also knew, if nothing else, that she would have to handle the issue of extricating from

---

[69] A good example of this is the ethics of campaigning on the stump. As Beerbohm (2016: 405) argues, politicians engage in "normative transactions" with their voters, serving an important function not only epistemically but in establishing the right agency relationships between voters and candidates. As an inherent feature of these systems, "the normative powers [are] liable to abuse" (ibid) in the form of lying to and manipulating voters. The challenge for Paul, then, is to learn how to grapple with these powers in a way that does not cross over into (unjustifiable) abuse. At a minimum, Paul is learning how to use tools that *can* be abused and exploited, even if he himself remains vigilantly reputable.

the unjust choices of her predecessors, which would itself entail moral costs and risks. She was volunteering, even in the absence of terrorist attacks, to take on a role for which she would ultimately be responsible for determinate moral wrongdoings.

There are analogous cases in the real world. George W. Bush, for example, became President with no expectation that he would soon have to navigate the acute ethical dilemmas of a post-9/11 world. Though he could not have anticipated the exact nature of the DE that would face his Presidency, it would have been political malpractice not to be prepared for DE to arise.[70] Questions of terrorism, foreign intervention, emergency management, dealing with hostile and autocratic regimes, and the like are interminable. The same is true for every political leader who found themselves grappling with difficult episodic decisions, such as whether or not to lockdown, in the face of the COVID-19 pandemic. Dealing with serious emergencies seems to be part and parcel of holding positions of great power. Hillary Clinton's famous 2008 advertisement, asking voters who they want answering a hypothetical 3AM phone call, appealed to that common wisdom: that to hold political office is to make oneself liable for emergency decisions.

These emergencies may not reach the level of the supreme emergency that Walzer speaks of, but averting extreme catastrophes from emerging, and being prepared to do the ATC-justified thing if they do, no matter how unsavoury it might be, is also part of the job. As one of their first official duties, British Prime Ministers are asked to write 'letters of last resort', instructing nuclear submarine crews on what to do if a nuclear attack has destroyed the British government. Prime Ministers may not take the giving of such instructions seriously, but in the event of disastrous moral luck or exceptionally poor statecraft, a consequential dirtying decision may already have been made. Usually writing such a letter constitutes a particularly

---

[70] Which is not to say that the Bush administration handled these issues well. The Iraq War and the Bush administration's use of torture, for example, both provide strong cautionary tales about an over-willingness to identify potential DE scenarios.

high-stakes DC (or, given your views on nuclear posturing, DR). However, if the letter becomes needed and its orders are acted upon, it can all of a sudden be assessed as an episodic choice, one of balancing evils. A DC choice thus becomes a candidate for DE. If their dirtying decision is justified under the circumstances, it will represent DE; otherwise, it will represent an act of gross wrongdoing and nothing more.

Part of the disposition entailed by DC is to be prepared for DE. The political leader must develop a "tough" disposition (Galston, 2005: 75-94), one that avoids "squeamishness" about tactics that are condemned by ordinary morality while avoiding an outright callousness. They must prioritise the traits associated with "experience" rather than "innocence" (Hampshire, 1989: 170), not merely because innocence is debilitating in quotidian politics, but because it can be paralysing in these highly sensitive episodic instances. This is not to suggest that the entirety of DC can be reduced to the propensity to face instances of DE. The dirtying dispositions and traits inflicted by DC go beyond those that can be justified by a discrete episodic analysis, and while what counts as DE will be changed by an agent's DC, the dirt of a DC itself is more than the sum of all the dirt that is made in discrete decisions that fall within an agent's DC.

Determining which dispositions and skills should be cultivated is a difficult balancing act. The danger can be found not just in innocence, but in the other extreme too. As well as callousness, there is a danger in being too eager to spot opportunities for DE, especially if one is drawn towards the dramatic figure of a tragic hero at the heart of "static" conceptions of DH. The agent must be prepared for cases where these moral violations are justified and thus must be willing to think the unthinkable. However, they must not see DE choices where they do not exist, or automatically treat each episodic choice as if it was unavoidable. This is one of the dangers associated with politics as a DC, and one of the ways in which the dirt arises. It can lead the agent to being on the precipice of committing moral violations *unjustly*, engaging in

motivated reasoning to justify such acts, or becoming guilty of acts that are simply evil and immoral. There is no plausible philosophical system that can prevent this entirely. However, being cognizant of the dirt and the costs of preparedness is helpful, as is the demarcation of dirt into these broader categories.

Another way that DC and DE interact is that DC helps us to understand the terms of reference for second-order dilemmas that arise because of discrete instances of DE. As discussed, the aftermath of many instances of DE creates temptations for further acts of wrongdoing, such as hypocrisy or secrecy around a moral transgression, instead of the admission and rehabilitation that the orthodox DH model proposes as necessary. How an agent should approach these second-order dilemmas will need to be assessed in accordance with the demands of their vocation, such as the self-justified goods of politics. Finally, DC can interact with DE in terms of the kind of practices that it makes possible. Certain actions in DE instances will be impossible if there is not an institutional practice already in place. Developing such practices can be either a DR decision (if its normalisation represents something pathological) or a DC decision (as would likely be the case with many intelligence and espionage practices).

3. The Limitations of Dirty Careers

Why can the DC not just encompass the entirety of political dirt? To enter politics is to run towards DE, increasing one's moral vulnerability towards them and taking on responsibility for their resolution. Similarly, accepting the self-justified norms of politics, including its contingent rules, is a key part of what it means to call politics a DC. Why, then, is the DC observation not sufficient, in terms of providing an explanatory, prescriptive, and normative description of what makes politics so morally distinct? In this section, I look at some of the limitations of the DC explanation. These are not intended as criticisms of the DC framework

per se. They instead highlight issues and concerns that the DC view cannot answer on its own, and thus show why the other facets of dirt need to be maintained.

The DC view is, in essence, a weaker version of the claim that politics represents entering a separate kind of moral existence. The stronger claim would be that the DC observation is grounds for deeming traditional DH problems as of limited or no importance, and that the dirt that arises from decisions we face *within* a way of life is no longer of primary interest. This stronger thesis, one that goes beyond viewing a DC as a source of dirt but as placing us entirely within a new moral landscape, is what I shall refer to as the Different Worlds (hereafter DW) approach. The DW approach holds that the answers to DE and DR questions are governed entirely by reference to the dirt of a career. The existence of a separate political morality purportedly answers all these questions, rather than merely conditioning said answers.

DC is a crucial observation for understanding why politics has the reputation it has and why some *pro tanto* wrong acts become part of a justified whole. My intention here is not to argue against the DC view or to put forward a sanguine view of politics wherein it is compatible with a conventionally morally harmonious lifestyle, but instead to argue against turning the DC view into a more robust DW view. DC is best seen as part of, rather than the entirety of, our understanding of dirt in politics.

Sometimes it can be unclear whether a thinker adopts the DW view or instead has more affinity with the less demanding DC claim. However, it is important for my purposes both to oppose the DW view and to caution those who agree with me on DC not to take this reasoning to its DW extremes. Walzer was very critical of the DW view, where political agents are subject only to "prudential" as opposed to moral forms of control (Walzer, 1973: 179) or where they acquiesce more totally to the "demon of politics" (Walzer, 1973: 177-178). The static nature of Walzer's framework has, as discussed, opened up the DH debate to further understandings that try and recentre these views of politics. The DW view can be found, most compellingly,

in Machiavellian and neo-Machiavellian understandings of the DH problem. The division between politics and ordinary morality is "intractable and perpetual", encompassing "exhaustive and incompatible" ways of life that cannot be reconciled (Tillyris, 2015: 64).[71] The politician "sinks deeper into the dreary arena of politics", becoming "accustomed" to the dirt that surrounds them rather than caught up in emotions such as tragic remorse and crippling self-doubt (Tillyris, 2015: 67). Political virtue is antithetical to ordinary virtue (Tillyris, 2015: 73), ordinary morality jeopardising political existence whenever it crosses that barrier. Thus politics requires a full-throttled and dynamic view of DH, instead of the focus on paradoxes of action, where political standards of excellence require the cultivation of a partial viciousness (Tillyris, 2016a). Moral purification as a goal in politics is neither possible nor desirable (Tillyris, 2018).

I agree with much of this analysis, including the unsuitability and undesirability of a focus on moral purification for politicians and the demarcation between political and ordinary integrity. What I want to emphasise is that there are limits to this: that these insights should be used to inform a DC view rather than becoming a DW view. There are tensions between ordinary morality and political morality, and many of these are intractable and thus inappropriate reasons for tragic remorse and the other emotions that are used in traditional DH analysis. The choice to enter politics does not licence all the viciousness that follows and does not capture all paradoxes of action that have interested DH theorists. Not all tensions between politics and morality are immutable or intractable, even if many of them are. There are limits— moral limits—to politics' capacity to be considered on its own terms. These are instances where

---

[71] Machiavelli's counterpart to political morality was more specifically Christian morality, and thus a stronger variant of "ordinary morality". We can, of course, operate with a weaker version of ordinary morality, such as the kind that Shklar talks about, where there are far fewer completely intractable incompatibilities between these ways of life. Adopting this weaker form of ordinary morality still allows, as DW views contend, that "there is no escape" (Hollis, 1982: 396) from the messiness of morality more generally. This makes the space between a political DC and ordinary morality easier to bridge. This is one way to defuse the allure of the DW claim while maintaining the positive attributes of the DC framework. I think this is important, but I choose to focus on negative arguments for this position in this chapter.

feelings such as tragic remorse and anguish might be more appropriate, where we do not want political actors to sink deeper and deeper into the depths of politics, and where we want the realities of political conduct to move a little bit closer towards a more acceptable baseline. Politics can be a practice and a way of life, and taking this choice can recontextualise everything else about an agent's moral universe and require certain vices to be cultivated, but there are limits to what this choice can explain that I want to emphasise.

In the remainder of this chapter, I discuss why I oppose the DW view and limit the application of DC. Firstly, looking at political morality *only* through the lens of DC reduces our ability to distinguish between politicians. Secondly, restricting our focus to questions of DC limits our ability to distinguish between the acts that politicians take. Thirdly, an exclusively DW focus leads to an overly narrow view of what constitutes a good politician. Fourthly, there are other constraints and impositions on a politician's moral life, which cannot be subsumed by the dirt of the career.

3.1 Distinguishing Between Politicians

We need a way of morally distinguishing between politicians and the political projects that they pursue. Reducing our analysis to DC obscures this. If we give primacy to the choice between different moral worlds and life-plans, assessments of conduct *within* that world are neglected. Tillyris (2019b) argues that suffering and remorse are inappropriate responses to the wrongdoing that politicians are asked to perform, as such acts of wrongdoing are part and parcel of political life. This seems accurate for the dirt that arises from choosing a particular career or world, because being racked with guilt in the face of the quotidian sacrifices necessary for day-to-day success might well be debilitating. However, there are choices that can be taken *within* this profession. Value-conflicts and norm-breaking discretionary choices that go beyond these

quotidian requirements will still exist. How politicians navigate these will vary. Furthermore, it would be a mistake to insulate assumptions about what is necessary from empirical or moral scrutiny.

Though there is plausibly dirt that comes from renouncing innocence-as-a-disposition, this would not capture all the dirt that might arise from DH problems. Cruelty is rarer in politics than the almost ubiquitous role created for hypocrisy; boundless cruelty is much harder to square with political goals such as stability and power (Tillyris, 2017: 489). It would be odd, then, to embrace a unifocal DW view that treats cruelty and hypocrisy as deserving the same conceptual treatment. Framing politics in this way would treat far too many unsavoury aspects of politics with the same brush, ignoring the many gradients of severity and extent.

In short, even after politicians have embraced their DC, this does not mean there are no longer important moral differences between them. These differences go beyond the efficacy with which they embrace their DC. DC allows for considerable hypocrisy in ways that do not extend to cruelty, but politicians may still be presented with DE where they must choose whether or not to authorise cruelty. These cases are not reducible to, nor their answers determined by, the choice they made to enter politics in the first place.

3.2 Distinguishing between acts

The supposed immoralities required by different categories of dirt possess different qualities and attributes and thus warrant different moral descriptions. DC helps capture this observation. However, making DC the whole story, as DW would, pushes against this insight. Something that is an intrinsic part of the vocation of politics should be treated differently from the result of some extraneous immoral imposition. In general, it can be justified to subsume certain practices under the general dirt of choosing a particular profession and adopting a broader

moral outlook, but other acts should retain the higher burden of moral proof (De Wijze, 2009: 317-318). This can be because of their moral gravity and the severity of what is at stake (normalising a heinous rights abuse, for example, is a more disturbing prospect than normalising grubby politicking), but also because some acts can be described as falling within political role moralities more comfortably. Others retain their dissonance with both general and political morality.

An example of this can be seen with compromise: some cases seem to be uncomplicatedly good, others more ambiguous, others still involving serious moral wrongdoing. In supporting his argument that politics requires abandoning an ethics-first approach, Tillyris (2017: 482) points to the ubiquity of compromise, treating it as something inherently morally dirtying, as "intertwined with betrayal, unfaithfulness and deception". Even if compromise represents something suboptimal from the perspective of our individual substantive moral commitments, saying that this therefore makes it always morally compromising is too strong. Sometimes compromises can be dirtying, but this is usually because there is something *pro tanto* wrong about the substantive content of the compromise itself, or that it generates a moral remainder. Much like the motive of political success, the act of compromise is not in itself immoral. Some compromises are benign. Some are even laudable, as they are important to democracy, show respect to others, or are praiseworthy for other reasons (Gutmann and Thompson, 2012). Compromise can serve as a positive good, as a remainder-free way of addressing some of the paradoxes and tensions of democratic politics (Bellamy, 2012).

Other compromises are certainly non-ideal, representing a regrettable and lamentable deviation from the good, but do not actually reach the level of doing wrong. Coady (2018: 2) cites the example of a party in coalition. The leader of the governing party appoints a competent politician from the other party to head a ministry over a more adept and better-liked colleague

from within his party in order to allow the coalition to endure. Though the leader is understandably less than pleased with this situation, and "certain valuable relations and ends are put under strain by it" (ibid), it is not in the same category as other types of wrongdoing. A politician may need a certain DC-adjacent disposition in order to do such a thing with the necessary sturdiness and resilience, but there is no dirt in the act itself. It is the disposition here that suggests dirt, and the cumulative impact of many quotidian compromises of this nature, rather than the decision itself.

Some compromises are, of course, wrongful. Some are paradigmatic cases of the DH dilemma. Compromising with immoral actors, helping them use immoral means to achieve immoral ends, is a wrong, and that wrongdoing is not annulled if we are acting with laudable intents or our decision is ATC-justified (Day, 1989: 481-485). If a compromise furthers the "evil projects" of other agents (de Wijze, 2007: 11-16) or is coerced as part of a complex of immorality, it certainly merits treatment as an instance of DE, or even DC or DR if sufficiently normalised. Even when the compromise with evil that we are making is ATC an easy one, we must still accept that we are helping some evil ends arise.

We can, quite easily, say that compromising is an essential part of being a politician. Even internally within a movement or party, compromise will occur. However, the moral character of such compromises, and thus the dispositions needed to make them, seems to vary. Some compromises seem to be good, a kind of virtue that is neither at odds with nor necessarily rare within private life. Others seem to be paradigmatic cases of the most severe instances of DH. Others occupy a middle ground, a dirtying yet ultimately quotidian feature of well-functioning politics (Bellamy, 2012: 447; 456).

Whether an act counts as dirtying can depend on the broader context. The boundaries between DC and DR can change over time too. What is seen at one point as an inevitable consequence of politics, an unavoidable necessary evil, can in the future be found to be

avoidable with the right institutional design or motivation. A good example of this is the existence of privileges accorded to warrior classes (Hampshire, 1999: 87), where what was once seen as an obvious reward (that is, non-dirtying) then began to be seen as undesirable yet unavoidable (or DC) and finally as an egregious and outdated injustice (a form of DH if it remained ATC-justified within a specific context). The attempt to secure private liberty by following the liberalism of fear is a useful example of how political reform can change what is seen as necessary, as such efforts reduce the scope for personal glories and vendettas to weigh in on government policy (Shklar, 1984: 214-216). It is also a useful example because it shows the limits to what can be achieved with institutional reform, as issues of personality and leadership still creep in regardless (Shklar, 1984: 220-221).

The ubiquitous nature of compromise within politics, then, cannot, in and of itself, describe the moral character of political action. Some compromises are of a benign character. Others are only dirty because of the dispositions required to engage in them habitually. Yet others are difficult decisions of the kind that merit DE analysis. To group the moral character of all compromises together is to neglect these differences. This is a mistake whether it is done to create an all-encompassing DW view, wherein the dirt of the compromising disposition describes everything, or to defend the orthodox DH view, wherein all dirtying compromises are assessed as episodic aberrations.

3.3 What Constitutes a Good Politician

The third criticism of the DW view is that it is too essentialist about what constitutes a good politician and underplays the importance of differing political skill sets. This concerns differences in scale, context, and level of politics. The DC of a local councillor is different from the DC of a US President, for instance. By saying politics is a wholly different world and that

entering this world captures the vast bulk of political dirt, these differences within politics are ignored.

A DW account also neglects the benefits that diverse political skillsets can bring to the political realm. For example, Dovi (2005: 129) argues that a morally healthy polity requires not only actors who compromise their integrity in pursuit of political goals, but those that refuse to do so. More generally, different politicians relate to their roles in different ways (Mendus, 2009: 118-120). Even if they operate within the same legal and socioeconomic structure, politicians may end up occupying the same roles for different reasons or for different ends. Those politicians bring different projects and dispositions to those roles, and rather than that diversity representing naivete or suboptimality, it can in fact be a strength. If role morality requires a complete detachment from who agents were before they adopted a particular role (Griffin, 1995: 264), these benefits can be lost.

This puts aside how difficult it might be to create an immutable and distinct space for a political role morality, or even a distinct class of political actions. If nothing else, what is a politician? The activist, the candidate, the backbench legislator, the head of government? These all have very different role moralities. Saying that they are bound by one and the same role morality, even within the same time and place, seems implausible. The DW view could say that there are multiple worlds, but given the interactions between these occupations, such a strict demarcation between worlds would be constantly under strain.

Similarly, not every politician can be said to have made a *choice* to embark along a different moral path. Politicians who think they may be able to play by conventional moral rules may well be naive in most cases, but sometimes a politician might find the rules that the realm of politics is governed by are degenerating around them. Alternatively, they may enter politics as a form of political 'self-defence', where political activity is the only foreseeable avenue to protect one's rights or core interests, or even be drawn into politics by a process akin

to sortition. Decisions that are part of the standard DC for most politicians can be encountered as an initial DE for others, as they are starting, legitimately so, from a different baseline.

This is further complicated by the fact that whether a realm counts as political is not a binary matter. This is a scalar question. For example, politics can be viewed as a tonal quality. Aspects that can properly be described as political can be found in many other realms of life and need not be consistently found within the political realm (Morgenthau, 1933 [2012]: 100-102). Office politics is a good example. Participating in office politics is presumably not a DC on a par with professional politics, but there is still a demand for dispositions like political virtues. For many, the need for political virtues ebbs and flows, its importance increasing and decreasing, with non-political demands and virtues retaining their own important roles.

3.4 Other Constraints on the Political Process

Finally, there will be causes of moral tension and conflict for politicians that are not part of the standard political process. These causes cannot and (perhaps more importantly) should not be viewed as something that is inherent to their political career. These fall outside of the domain of DC, as being outside what the self-justified politics can and do justify. These can take two forms: firstly, extraneous or irregular disruptions to what is usually accepted from a political career. Secondly, those features of a political career that can be changed.

This is where the DW view is vulnerable to critiques around decontextualisation, namely of the variety Coady spoke of. For example, certain acts such as repression are themselves constraints on the political process (Luban, 1980: 397), suppressing its ordinary operation or at the very least departing from what politics should be. Similarly, civil war is also arguably a failure of (democratic) politics. If they *seem* to be part of the DC of politics, then they go beyond the standard DC. They represent a different baseline and should be challenged

(whether as DE or DR) rather than treated as inevitable. A general distinction between acceptable and unacceptable regimes (Galston, 2010: 408), and the respective conflicts that they bring up, is one worth maintaining.

Some of the causes of DH dilemmas cannot be described as ordinary politics. For example, agents might find themselves placed within a "complex of immorality" (De Wijze, 2009: 309) or forced to grapple with the "evil projects" of others (De Wijze, 2005: 456). These should rightly be viewed as breaking, if not quite the harmony, at least the rightful baseline of political conduct, as going beyond what people were 'signing up for'. Similarly, the actions done in response to these acts go beyond the level of dirt that political actors are typically asked to take or asked to normalise. These can be acts that, for example, go against the letter and spirit of prevailing political norms (De Wijze, 2009: 313), whether in a 'lone wolf'-style individualist sense or some more enduring manner, or that violate "legal archetypes" (Waldron, 2005). These rightly entail guilt and restraint beyond that which politicians should normally bear and what they take as just part of the profession.

The language of political necessity, as created by the self-justified goods and ends of politics, is a powerful one that politicians must grapple with. Still, it is a language that can be changed. What counts as necessity should not be treated as an immutable fact (Coady, 2008: 89; McDonald, 2000: 188-189). Where purism struggles is that efforts to change the language of political necessity must themselves engage with said language. This is especially true when talking about conduct itself. Moral judgements must be "intelligible" to those who are engaged in a political environment if we want said judgements to translate into meaningful commentary (Philp, 2007: 74). However, while this contextualisation is important, it also should not be a blank cheque. Appeals to necessity, and the relative fortune or misfortune an agent finds themselves dealing with, must not be used to completely shake off any notion of personal responsibility (Shklar, 1984: 167).

4. Conclusion

Politics as a set of practices invites a certain degree of viciousness and a willingness to routinely engage in acts such as the use of violence and the compromising of key moral rights that would otherwise be unpalatable. Politicians must abide by the self-justified rules of politics and accept its normative standards and valued dispositions. This is an important observation, both for what it tells us about the nature of politics and as a necessary refinement of the Dirty Hands claim. The idea that politicians can, or should, try and cleanse their hands after a DH action is an oversimplification. The idea that their hands were clean before the action in question is usually outright mistaken. By treating issues of political dirt as solely a matter of actions, DH thinkers have often bought into an overly static and decontextualised view of the phenomenon. I have also defended the underlying metaethical claim that moral *perfection* within the political sphere is unattainable and that compromise and conflict are intrinsic components of politics (Tillyris, 2013: 251-253), which entails that there is a perpetual disharmony between politics and the traditional precepts of morality.[72]

Even if we were to construct an ideal politics that sidesteps all surmountable empirical limitations, seeking an absolute harmony between politics and morality would be folly. Such a quest "mischaracterizes the qualities conducive to virtuous political practice, the nature of political integrity as well as the peculiarities of political relationships and the messy context in which politicians operate" (Tillyris, 2013: 254). Good political practice requires cultivating some dispositions and eschewing some virtues that would otherwise be valuable. Ultimately

---

[72] I do not think agreement with Tillyris on this point, however, is essential to my argument. If one takes the view that a morally perfect politics is 'merely' empirically impossible or implausible, rather than a deep conceptual impossibility, distinguishing between the dirt in Dirty Careers and other types of dirt will still be, I believe, a worthy endeavour.

the main uses for DC analysis are helping define the character of politicians, highlighting the quotidian everyday cases of dirt (as distinguished from those that represent a more serious moral problem, both qualitatively and quantitatively), and working out where politics requires a deviation from ordinary morality and where it does not. Trying to get someone fired is, for example, par for the course for politicians. Such is the essence of electoral campaigns, yet it requires a change in thinking beyond what ordinary private life usually asks, and that has implications for the dispositions held. Ordering torture, however, is not.

There is still, therefore, a space for understanding other types of political dirt, and that this calls for a—revised and limited, admittedly—space for the more conventional forms of DH analysis. The DC thesis is correct. Taking this further and applying it as a DW answer that encapsulates all questions of political dirt, however, would be mistaken. The starker and more shocking political conflicts still require some great attention to be paid to paradoxes of action, and there are limits to what the self-justifying norms of politics can be expected to excuse in terms of individual behaviour. This is why I have drawn the distinction that I have between the stronger DW claim and the DC claim.

Certain acts and practices can be required by politicians which generate moral dirt that goes beyond the ordinary taint of adopting the political lifestyle. Some of these practices can define a career, and the unfortunate necessity for such practices can be very difficult to overcome. However, the fact that these unsavoury practices *can* feasibly be changed or altered, or that they inflict particular kinds of especially intense moral dirt, means that they cannot just be subsumed into the DW claim. The DC claim is part of a coherent account of political dirt. The DW claim succumbs to many limitations, including similar issues with decontextualisation that plague the orthodox DH view it seeks to debunk.

CHAPTER VI: DIRTY RULES DILEMMAS

Throughout this thesis, I have defended the notion that politics can, and does, entail accruing ATC-justified dirt. This is not a novel view in the literature, of course. It has been proposed both by DH theorists focusing on discrete choices that inflict determinate guilt and those who argue that entering politics requires subscribing to a whole new set of standards. Neither of these views on their own explains the whole story, however. Instead, I have incorporated these views into my own framework. Dirty Episodes and Dirty Careers represent different manifestations of the same core idea of ATC-justified dirt. As a result, they are best understood as complementary rather than adversarial ideas. Still, the tripartite model of dirt is incomplete thus far: there is another way in which circumstances can render a dirtying course of action permissible.

I now turn to the most analytically novel category of dirt proposed in this thesis. Dirty Rules (hereafter DR) describes those structures, institutions, and norms which unnecessarily and unjustifiably permit, incentivise, or otherwise legitimise corrupt or corrupting behaviour. These institutions create dirt that falls not only on the institution and their authors, but on those who use these practices and play in accordance with these dirty rules. DR dilemmas arise when it is ATC-justified to engage with these institutions and incur the associated moral pollution. The resulting dirt can arise as part of a specific decision, as part of an overall practice, or manifest itself in the skills and dispositions an agent chooses to nurture.

In this chapter, I will firstly sketch out why I think accounts of dirt need a space for DR theorising. Once this has been established, I will discuss what this concept covers: its necessary conditions, the examples it best explains (particularly Edna's and the North Takoma gerrymander), and how it interacts with other concepts. Next, I will discuss what creates the dirt in DR situations. Finally, I discuss the phenomenology of this dirt and some of its

implications: in short, the justifiable usage of DR acts necessitates a greater degree of proportionality than more benign or DC political practices, and that while the acts should be viewed with tragic remorse, democratic agent-regret is appropriate for the system itself.

1. The Need for DR Theorising

Of the five examples introduced in the introductory chapter, two in particular call for DR theorising. Firstly, Edna, a politician whose country has a deeply corrupting campaign finance system that is so egregiously unjust that she views it as a kind of legalised bribery. The moral wrongdoing of bribery is not diluted by its legal status. Edna is passionately opposed not just to the existence of this system, but to the acts that it serves to normalise. Edna's first preference would be, if possible, to run for office without engaging in this system. She would be willing to accept some kind of competitive disadvantage. However, she makes a reasonable and well-informed assessment that political success *requires* engaging with the system, and therefore requires soliciting legalised bribes. Otherwise, her campaign will go nowhere, and perhaps divert resources and votes away from other candidates that share her substantive goals but can advocate for them more effectively. The price of political competition, therefore, is to engage with this unjust system and the unjust practices that it requires.

Edna has already accepted that politics is a Dirty Career, with its own rules and virtues, but she does not think a system *this* egregiously flawed should be accepted as another part of politics. Suppose a confidant offers her that familiar refrain "if you can't stand the heat, get out of the kitchen". Edna would accept there is some truth to this. She is willing to (metaphorically) break into a sweat, however Edna thinks the system is demanding more of her than that. The "kitchen" is far hotter than it needs to be to meet its functions, so to speak. This unnecessary heat both harms the practitioners and frustrates the achievement of the kitchen's goals. Turning

the heat down (reducing the immorality of the campaign finance system in Edna's case) requires working within the "kitchen" as it currently exists. That is the frustration for Edna: she rightfully wants to make the system fairer but must work within the system as it currently exists to do so.[73] She might even find that, in order to obtain power, she must contribute to the processes that are turning the heat up.

The second relevant example placed us in the position of legislators in the state of North Takoma. The only way to counter the gerrymander of neighbouring South Takoma, and restore competitiveness to federal congressional control, is to gerrymander in retaliation. The North Takoman legislators are well-aware of the wrongs of gerrymandering, but not to gerrymander in this case would leave the outcomes of federal elections a foregone conclusion. I have argued elsewhere that to gerrymander in this situation would be to engage in permissible wrongdoing (Sarra, 2022). This problem thus belongs within the DH framework. However, I do not think it can be understood as DE. Granted, systematic gerrymandering is usually such a severe wrong that it necessitates a lesser evil justification of the kind that is found within DE conflicts: its determinate wrongs can usually only be justified when it is so obviously a lesser evil. But that is not what is at stake in the North Takoma situation. What makes gerrymandering necessary, and ATC-justifiable, is that it is now a necessity to achieve fair democratic competition.

Even a DE structure that is sufficiently attentive to contextual costs and requirements will be an inadequate way for understanding these cases. This is partly because the moral stakes are different. Usually, DE cases arise in the pursuit of a lesser evil. This terminology is too dramatic and the threshold too high for these cases. What is at stake in these cases is not averting some greater wrongdoing, but instead the ability to compete within an imperfect democratic system. Edna and the North Takoman legislators might well have morally important

---

[73] This will not always be the case, of course. Edna might find that the system can be effectively challenged from the outside, or alternative arrangements could be set up. Assuming an alternative of that kind is feasible and accessible, the dirt that comes from engaging with the system as it currently stands will increase, even if it remains an ATC-permissible course of action.

goals and considerations, and be seeking to avert morally undesirable outcomes, but their predicament can be understood even without substantiating these. Perhaps these cases could be restructured to have a great evil at stake, but this is not necessary to draw out the moral quandaries in these situations. Political considerations are behind them both.

DE cases necessarily involve a determinate wrong that is to be prevented, rather than the acts being justified by the general pursuit of the goods of politics. DE cases have some discrete inciting cause: the demands of a despot, the needs of a military strategy against an appalling foe, a tricky political situation borne from the complexities surrounding a specific issue or policy. There is something that specifically makes the wrongful behaviour justified beyond merely the 'demands of politics' or the open-ended good of political participation. Something has gone wrong. Other agents have inflicted bad moral luck upon the agent who must decide.

In contrast, DR dilemmas arise from the institutional structures that we must engage with to achieve democratic change.[74] They are structural.[75] A successful political agent will have some responsibility for these structures, and an aspiring political agent will end up with some responsibility for them. They are also, however, remediable and contingent. DR cases represents aspects of our political framework that political actors ought to change and indeed have a duty to change. On top of the duties to change these practices, there are moral reasons to minimise their usage. This means that the dirt entailed by any given DR dilemma should not be regarded as an intrinsic and inevitable part of politics. This is what distinguishes them from DC.

---

[74] I do not believe that DR are unique to democracies, but I focus on how they materialise within democracies here. In non-democratic contexts, they are likely to be even more prevalent and even more embedded, given how autocracies reduce opportunities for morally ideal methods of achieving political change and that using the mechanisms of autocratic government is liable to be dirtying in a way far greater than democratic institutions are. So long as a move away from autocracy is morally accessible, these methods will count as DR.

[75] As to the definition of "structural", I believe aspects of Young's (2011: 52-64) definition of structural injustice can be adapted here: social rules and positions, "produced through the actions of individuals" (McKeown, 2021: 3), are objectively constraining in a way that can have unintended consequences.

1.1 When do DR Dilemmas Emerge?

We first turn to what constitutes a DR scenario. For a DR scenario to exist, an unjust structure is necessary but not sufficient. On top of this, there must be something wrongful about engaging with and utilising the tools provided by the structure. There is wrongdoing involved in acting as if acceptable standards were not in place. The agent shares in the moral responsibility for these unjust structures and their impacts. The agent cannot simply pass all responsibility onto the institutions or the institution's authors.

In a nutshell, the harm in DR arises in two steps: firstly, there is a structural flaw. These flawed institutions create openings for wrongdoing and require deviations from morally acceptable rules. Secondly, it is the discrete acts of using those institutions that are constitutive of the wrongdoing—or at the very least make up a large part of it.[76]

The institutional existence of this flaw is not in itself sufficient to create DR scenarios. Suppose there is a flaw, but political actors are not exploiting it. This could be out of moral conviction, an absence of sufficient incentive, timidity in their ability to use the tools effectively or to their advantage, or ignorance of their ability to exploit these tools. In such cases, this flaw should be fixed as a matter of prudent institutional design, but it is not (yet) a DR scenario. That "others [have] set the terms of the competition" (Walzer, 1973: 163) is the core of a DR scenario. This term-setting will not solely be via legal structures or publicly espoused norms but will instead be authoritatively revealed in what becomes necessary to achieve some measure of political success.

---

[76] Like with DC, DR is more than merely a greater propensity to face DE instances or an aggregation of every DE instance that a political situation might create. Instead, DR covers dispositions, diurnal acts, and habits of the kind that elude episodic understanding. DR scenarios entail bargaining with the corrupt 'rules of the game' in a quotidian way that defies the 'lesser evil' standard of justification attached to instances of DE.

For example, a system where politicians have control over the redistricting process in a way that means they *could* gerrymander is a flawed system, even if they do not use their power in this way. We will likely have reason to want to change this system, even in the absence of observable gerrymanders (Issacharoff, 2002: 643-645). Politicians have the means and method to gerrymander. This is not a DR scenario just yet, however, for no gerrymandering is yet taking place. For gerrymandering to actually occur, this power must be used in a certain way, with politicians choosing their voters to achieve certain ends (Beitz, 2019: 7).

Only when politicians actively gerrymander do the distortive and damaging impacts of gerrymandering materialise and voters actively begin getting harmed by gerrymandering. Thus emerges the incentive to respond with retaliatory gerrymanders, to balance out the aggregate impacts on our decision-making processes (Sarra, 2022). Only once this point has been reached, as it has for the North Takoman legislators, do we find ourselves in the territory of DR. For gerrymandering to be justified prior to this, the more demanding threshold of DE must be met. Analogous reasoning could arise in a system with the as-yet-unrealised potential for campaign finance abuses to emerge. If that potential has not actualised, actors like Edna face no DR dilemma. However, there are still *prima facie* reasons to change these imperfect systems to prevent them from being abused in the future. As is, however, abusing the campaign finance system cannot be justified by pointing to the terms of the competition.

The moral wrongdoing of utilising DR practices, and the corollary moral ought to abstain from them, is not alleviated by the fact that institutions and structures require or endorse these practices. The wrongful nature of these acts is not allayed by societal condonement or political necessity. The discrete harms inflicted by that structure are brought into being, or at the very least exacerbated, by how politicians actually use them and make them part of the terms of competition. The wrongs in DR dilemmas are *pro tanto*, not merely *prima facie*, much

like other DH wrongs, and the structural nature of these wrongs does not render them defeasible any more than their being ATC-justified does.

DR can be seen as part of the political role morality, yet it is a different part to that occupied by DC. A role morality system can distinguish "between wrongs that a reasonably just[...]system permits to be worked by its rules and wrongs which the [agents] personally commits" (Fried, 1976: 1084). DC concerns the dirt that falls on the personal agent because of the wrongs reasonably permitted by a decently just system. DR instead concerns the dirt that an agent personally commits but that the system wrongfully permits or requires, even if the system is reasonably just as a whole.

1.2 How are DR Dilemmas Created?

Paradigmatic DE cases have an identifiable and determinate antagonist: an actor or actors whose wrongful behaviour has placed the 'protagonist' in a situation where they can 'do wrong to do right'. DE choices can arise regardless of the protagonist's broader moral harmony or moral purity, and regardless of whether the protagonist has any responsibility for the broader conditions that generated this episodic issue. DE scenarios, in short, involve antagonistic figures who compel the protagonist to err from morality, beyond the demands of the protagonist's DC, in order to mitigate or challenge an extraneous immorality. In the extreme, these are terrorist figures and genocidal warlords, but they do not need to be so obviously contemptible figures. DE antagonists can be politicians who have acted corruptly or dishonestly, voters clamouring for an immoral policy, or past government ministers implementing an unjust policy that will tie the hands of their successors.

Paradigmatic DR dilemmas, in contrast, do not have such a cause. They arise from the norms, structures, and institutional frameworks governing political success within a given

polity. There need not be an identifiable blameworthy agent, nor is there an extraneous immorality that causes determinate moral conflicts. Instead, DR dilemmas and their associated dirt stem from remediable and structural flaws. A given political entity has structured itself in a flawed way and this is what renders dirt ATC-justifiable, as opposed to this being done by some inflexible aspect of politics or some risk of extraneous immorality.

The path to eliminating a given DR dilemma, from the standpoint of institutional design, is a clear one: change the institutions. Other polities have likely shown that this is possible. With DC and DE dilemmas, we can reduce their occurrence, certainly, but no institutional arrangements can change the fundamental rules of politics or totally guard against malicious actors. Any contingent institutional fact is open to change, however, and so a politics with any given source of DR problems being removed is, in theory, attainable. Moreover, there is dirt in leaving DR institutions in place. Wanting to change these dirty rules is a legitimate task of politics, rather than an admission of a disqualifying naivete.

1.2.1 How are DR Institutions Created?

How, then, do the institutions that facilitate DR dilemmas emerge? Not all DR institutions are created knowingly or culpably. Sometimes they are, of course. Take the institutions that govern politics in the United States. These institutions do not reflect coherent institutional designs, but instead grubby political compromises and the "felt 'necessary evils'" that these compromises required (Levinson, 2020: 386). Many of the founders, for example, viewed institutions such as the Senate as fundamentally at odds with their principles. Yet it was a price they were willing to pay in order to create the nation. They thus created institutions that at least some of the authors knew were flawed. Creating DR institutions, and subjecting other (future) political

actors to DR institutions and structures, can thus be the 'dirt' involved in a DE dilemma.[77] DR institutions can thus be created deliberately, and sometimes this creation can even be ATC-justified.

More interestingly, though, is that DR institutions can also be created by what Levinson (2020: 386) terms "honest mistakes". These mistakes can become apparent over time and can reflect an inability to futureproof one's ideas. In some cases, once laudable institutional innovations are rendered obsolete by changes in events (Levinson, 2020: 383-384) yet are already rigidly entrenched by this point. Kang (2020: 1426-1445), for example, argues that "mismatched" legal frameworks, created for a less partisan political climate, produce the current legal openings for gerrymandering within the United States.

To pursue the best institutional practices of today, perhaps even to rectify the DR flaws of today, might create new DR problems a few decades down the line. DR practices can thus arise not only from malice or neglect or compromise, but as an inadvertent and unpredictable consequence of any attempt to provide rulesets to govern political life. No institutional arrangement can completely guarantee against future developments such as the weakening of norms or the escalation of political competition. Institutional design thus has, as an inherent risk, the potential to one day create DR institutions as part of its DC.

As well as changing circumstances, value judgements can change too. Stuart Hampshire noted that moral conflict is an inevitable outcome of humanity's imaginative tendencies (1999: 40). For Hampshire, conflict is inevitable and underpins human pluralism (Edyvane, 2019). A final set of values will never be reached as a result. New ideas of "substantial justice" will continue to emerge, meriting new compromises and trade-offs, with these new conceptions

---

[77] This is another instance in which contextual sensitivity is paramount. Allocating this historical responsibility, however, may not be all that useful in allocating responsibility or liability in the here-and-now.

serving to "repeatedly widen the debate and open up cases of injustice that had hitherto been beyond the range of discussion" (Hampshire, 1996: 162).[78]

This means that rulesets once seen as acceptable may one day lose their acceptability, potentially creating DR dilemmas.[79] New DR openings arise as our views on justice change and evolve. Privileges that were once seen as necessary come to be seen as unjust (Hampshire, 1999: 87); natural misfortunes can become conscious political choices if maintained once the capacity to correct them arises (Hampshire, 1999: 84). We may not escape DR *permanently*, but each given instance of DR, though flowing from an inexorable part of politics, is a remediable and contingent structural feature. Preventing DR might well be akin to a game of whack-a-mole, with constant vigilance required. This is merely another way in which the inescapability of conflict in politics manifests.

Another potential place where DC can create DR is adversarial ethics. Adversarial ethics arise when the benefits of competition legitimise or incentivise behaviour that would otherwise be immoral (Bagg and Tranvik, 2019: 975-976). Adversarial ethics ground some DC practices, but do not legitimise the exploitation of all "socially destructive" practices in pursuit of advantage. Ethical agents encountered with such destructive opportunities must not merely abstain but work to pre-empt their adoption by others. They can do this by raising the costs of using such practices and putting up barriers. This is a moral duty, but also one of self-interest,

---

[78] As well as new conceptions of the good emerging (Hampshire, 1999: 38), the "current limits of political agency" will be pushed forward too (Hampshire, 1999: 48). Misfortunes that were once accepted as natural or falling outside the domain of politics may one day be understood to fall within the proper scope of political change. Institutions that assumed these misfortunes fell outside the bounds of political change may thus turn into DR, despite once being considered DC or not dirtying at all.

[79] How much agreement must there be that a particular ruleset is dirtying before we say that a DR institution has emerged? Whose perspective, in other words, is important for saying what constitutes DR? One way to address this is to look at societal agreement. Another is moral truth. Both Edna and the North Takoman legislators can make a strong case that what they object to should be decried by all who share basic democratic commitments. Another way is more subjective: to focus on the agent facing the situation and how they themselves view their choice. DR arises based on what the agent considers dirtying. While I think this last approach is the least accurate in actually identifying moral dirt, it will still be a useful framework for understanding how an agent (even one who incorrectly identifies dirt) understands their predicament phenomenologically and navigates their own issues of integrity.

to ensure that they are not hurt by other less scrupulous agents taking advantage of these destructive options.

Say it is too late for prevention, however. Perhaps the incentives were too high, or an immoral agent spotted the opportunity to gain an advantage via a socially-destructive opportunity first. These socially destructive practices have become normalised as part of the competition already. Not only can this create Structural Dirty Hands problems (to be discussed later) of an extrication kind, but it can also create DR predicaments when these adversarial ethics come to define a morally necessary role morality.

1.2.2 Hollis's Challenge

Martin Hollis (1982: 396-397) rejects the notion that DH can be solved by correcting defective institutions. Hollis notes the allure of this idea: by correcting defective institutions we could bring to justice the evil actors and evil projects that generate so many DH dilemmas. This neglects, Hollis contends, that the real source of DH is found in the inherent pluralism of politics and the compromises this requires. The state relies on "covert violence", hypocrisies, and similarly grubby ends, and from this "there is no escape" (Hollis, 1982: 396). I have already agreed with Hollis here to some extent: there is no escape from DC except to vacate politics altogether, and no way to completely avoid the risks of being confronted with DE or generating new DR institutions.

Where Hollis errs is in neglecting the possibility that some of these dilemmas can be solved. Dirt in politics is indeed the "nature of the game" (Hollis, 1982: 397). This does not mean that every instance of dirt in politics should be viewed as such. In particular, some of this dirt will stem from those institutions that we could remedy. These remediable instances of dirt should not be viewed with such defeatism. If politics unavoidably creates scenarios where we

must choose between, as Hollis (ibid) puts it, "the People at one end and the victims at the other", this is certainly tragic, but if we find out that some of these dilemmas, even if only a small portion of them, can be fixed by institutional reform, it would be remiss not to attempt to do so.

1.3 The Necessary Conditions of DR Dilemmas

What are the necessary conditions of a DR dilemma? Firstly, it is worth analysing the necessary conditions for a DR *institution*. These conditions are:

1. A serious flaw in the institutional structure of a given political framework. This flaw incentivises, legitimises, or requires wrongful behaviour.

2. This flaw is remediable. It represents something that can, and should, be fixed. It is neither a necessary nor a desirable part of a political system.

3. There is a moral imperative to change this system. There is a moral wrong entailed in leaving it in place.

4. It is morally dirtying to follow the standards of behaviour that this system promotes. Using these institutions inflicts moral dirt upon the agent.

The focus of this thesis is, of course, on ATC-justified or permissible dirt. When does the moral dirt that comes from engaging with such institutions become morally permissible? Holding the *severity* of dirt equal, the way in which dirt becomes permissible impacts its qualitative nature and thus the dirt's resulting impact on agents' character. DE represents aberrations and the intrusions of external immoralities. DC represents the uncomfortable but ATC-acceptable costs of a certain lifestyle. DR, structurally, lies in the middle between these two poles. Generally, DR acts are easier to justify than DE, and require less of a moral burden of proof to be met. However, their contingent nature means that they should not be normalised.

The necessary conditions of DR dilemmas are therefore as follows:

1. Political success requires engaging with DR institutions, and attempting to pursue political success within the system is ATC-justified.

2. The use of these practices leaves a moral remainder on the agent, and the agent should not use these practices as if they were morally benign political practices. There is a moral good associated with acting as if cleaner rules are not in place.

3. While these institutions could be changed, reform endeavours are by no means easy. Such change, as well as the pursuit of other worthwhile goals in the meantime, requires political success and therefore engagement using the rules as they currently exist.

DR covers both paradoxes of action and paradoxes of character. As a result, for moral agents, the necessary engagement with DR practices should be conducted with moral imagination as to how these costs can be avoided or mitigated, with proportionality in how these techniques are wielded, and with some attempt to preserve their personal integrity even as contingent political realities make full adherence impossible.

Note the assumption that attempting to pursue political success is justified. I will not defend that point in detail here. My arguments against purist abstention from politics have already been made in Chapter 2. Furthermore, I believe it is fair to assume that, within democracies, the right to engage in political activity is permissible for most agents.[80] This right is not destroyed or invalidated by the existence of DR institutions. Otherwise, that would reduce it to a "right-only-when-justified", to borrow Cullity's (2007) phrasing. Rights to political activity do not disappear when the political institutions are corrupt. Indeed, this is arguably when the presence of the right is at its most important, as it provides the means for making a better political climate possible.

---

[80] Even if one does want to attach other conditions such as reasonableness, baseline democratic commitments, or abiding by the law.

A more interesting way in which the first condition cannot be met is if the system is so corrupted, so broken, that it is no longer ATC-preferable to work *within the system*. Political success remains justified, but not success within this particular political system. Instead, working outside the system, or seeking to tear down or destroy the DR institutions, becomes more justifiable. I accept this possibility, though do not explore it in detail here. I believe it is a greater probability for non-democratic societies, which is beyond my present focus.

2. Where the Dirt in Dirty Rules Comes From

Unlike DE acts, DR acts need not be justified episodically, with each individual decision justified on its own as a lesser wrongdoing. Of course, the wrongdoing entailed by DR may well be of a severe enough quality that it can *usually* ground DE instances and require such lesser evil justifications. The democratic violations involved in gerrymandering seem to rise to this level: the abstract goal of political competitiveness is usually not enough to justify an episodic violation of a severe moral constraint. The acts that are permitted under DR contexts are usually those that would normally merit DE consideration. DR contexts, however, ask politicians to treat them as if they were more akin to being part of a DC.

Both Edna and the North Takoman legislators are trying to achieve worthwhile and respectable political projects by using the rules of democratic politics as they find them. The rules just so happen to give licence and incentive to engage in practices that, absent these structural features, would be unacceptable. All political actors operating at the same level must deal with these rules, regardless of the political projects they seek to deliver. The reasons, then, are agent-relative, those reasons that motivate ordinary political action rather than the extreme agent-neutral causes that tend to motivate DE instances.

DR dilemmas are not caused by specific interventions or disruptions to what might otherwise have been a well-functioning polity. They are not isolated incidents. That the polity's functioning is flawed *is* the source. Others have, in Walzer's (1973: 163) words, "set the terms of the competition". In these cases, it does not mean, as in DE scenarios, that the evil projects and unjust schemes of others have forced us into a scenario where countering them requires advancing or partaking in evil or unjust acts. It means that the terms of the competition, no matter how grand or trivial the stakes are, must be carried out in accordance with certain norms. Both Edna and North Takoma's legislators can correctly point to the corruption of their political climate as reasons for their behaviour, but the fact that broader society shares in the blame does not absolve the agent of any and all resulting moral responsibility (Shklar, 1984: 146-148).

What is the nature of the dirt involved in DR, then? Why is it that an agent does wrong when they behave as if cleaner rules were not in place and when they instead accept the contingent self-justified rules of politics? Without being able to answer these questions, it would be tempting to place DR entirely at the level of institutional critique, or to just accept it as part of the inherent dirt within the career of politics. My answer to these questions is to focus on the moral good that comes from rising above DR rules, and the corollary moral dirt that comes from sinking to their level.

2.1 A Duty to Behave by Cleaner Rules?

DR institutions are so dirty, so corrupting, that not only is there a compelling moral duty to change such institutions, but a moral good in behaving as if cleaner rules were in place. Edna would be morally better off to abstain from corrupt deals; to gerrymander is worse than to not gerrymander. These moral judgements remain even when the institutions incentivise these

behaviours. There is dirt in behaving in accordance with these dirty rules, and a corollary *prima facie* duty to behave as if cleaner (even if not perfect) rules were in place instead. What grounds this duty in democratic societies?

Part of this answer will be found in the specific harms or wrongdoing caused by the acts in question. DR structures will vary considerably in what wrongful acts they condone or facilitate. The wrongfulness might materialise in harm to individuals, in the corruption of character, or in broader structural harms. With the North Takoma gerrymander, the harms are in the form of both compromised outputs that unfairly burden a community's chances at achieving democratic success and in compromised inputs that do disrespect to democratic ideals (Beitz, 2019). These harms materialise only when political actors take advantage of the space for gerrymandering created by the system and apply regardless of how 'sympathetic' their motivations are. The dirt of gerrymandering thus belongs to the actors, rather than solely falling upon the system itself. The actors become co-authors of this continued narrative of wrongdoing, even if their individual contributions are justified.

Beyond specific contextual harms, however, there is also a duty to not comply with dirtying rules. Agents have a duty to rise above them (perhaps not to the levels of moral purity or ordinary morality, but at least something close to the levels of DC). What is the nature of this duty, and why is there personal responsibility for a collective and institutional failure? It is quite easy to formulate a theory as to why agents should have a duty to, where possible, remedy corrupt and dirtying institutions, and why they should value effective institutions where it exists. We can do this by pointing to the merits of better institutions. Identifying this value, however, does not immediately explain what duties arise for individuals in terms of how they interact with these dirtying institutions in the interim.

I turn to Janna Thompson's (2007) argument for why there is a moral responsibility to engage in political activism, even where individual contributions are highly unlikely to have a

noticeable effect. Democratic institutions allow citizens to work collectively to make systemic changes that "prevent or eliminate injustices", rather than just mitigating or countering specific cases as they arise (Thompson, 2007: 159). This explains why effective and fair democratic institutions should be valued by the citizens who live within those societies, and plausibly why the citizenry would want to shape those institutions into more just or effective forms.

What this does not do, Thompson points out, is justify why citizens have a proactive individual duty to use collective democratic institutions over other methods. Why should citizens not act individually and focus more on addressing specific cases on a one-by-one basis? The end goal might be less significant, but their individual contribution is more discernible. Similarly, this observation alone does not explain any specific duty about how to interact with institutions that fall short of that moral threshold or any reason why there is something morally valuable about acting as if better institutions were in place. Both problems are, at the core, a problem of many hands (Thompson, 2007: 156): a rational assessment of the efficacy of our individual contributions suggests a lessening of moral responsibilities that can, at the broader societal level, lead to passivity.

2.1.1 Democratic Duties

How Thompson solves her problem, and an approach that I think is fruitful for mine, is a shared commitment to making democracy work (Thompson, 2007: 163-166). The joint project of democracy, and its associated benefits such as accountability, create duties that exist independently of any assessment about the efficacy or lack thereof of individual contributions. For example, there is a commitment to accepting our losses.[81] This is why, Thompson (2007:

---

[81] Abiding by such losses, even when we regard them as morally wrong, can ground dirt too, particularly of the DC variety. See Waldron, 2018.

166-167) argues, we have individual responsibilities to campaign against unjust policies or leaders in our democracies. If we had the opportunity to do so but did not, then we are culpable for the injustice, in a way that an unsuccessful opponent is not. 'We', in this context, can refer to democratic citizenry as a whole. This obligation seems most pressing in the case of politicians, however. They have greater influence to change such systems and greater complicity and culpability with the wrongs of such systems.

We can also point to natural duties of justice (Howard, 2019: 178-181), and how these duties in turn ground broader democratic duties. Under this view, we have natural reasons to maintain and defend just institutions where they do exist; to reform them when they are unjust; and to create just institutions where they are absent.[82] Democracy is thus the default institutional framework mandated by this natural duty, as sharing the burdens of the duties of justice best shows recognition-respect to other actors. The preference for democratic institutions created by this natural duty is not necessarily absolute. "A deviation from the democratic default rule is justified if it is beyond reasonable dispute that such a deviation would advance the achievement of justice" (Howard, 2019: 189). But what if the institutions incentivise breaks from these democratic default rules that are not justified by this standard? This is where DR conflicts emerge, and the duty to abide by democratic default rules is broken (permissibly from the standpoint of the agent, impermissibly from the perspective of institutional design). In cases where the default presumption for democratic rule *should* hold, and there are resulting duties of democratic compliance placed on individuals, institutions that legitimise behaviour that take us away from these duties are dirtying. They inflict dirt by denying the recognition-respect that would come from fairly distributing the burdens of the natural duty of justice.

---

[82] Again, this duty is stronger for those with greater political responsibilities but can plausibly be applied to the broader citizenry as well.

Inherent within many acts of DR are discrete democratic wrongdoings that seem to transgress whatever duties actors might have to make democracy work. As well as duties, we could look at rights as well. Individual democratic rights (Brettschneider, 2007) that underpin a democratic right to self-authorship and self-ownership might be compromised by the discrete uses of DR structures. For example, gerrymandering can be critiqued as such a transgression of democratic rights (Sarra, 2022). Retaliatory gerrymanders hurt individual voters' democratic rights in the same way as unjustified gerrymanders do. A strict adherence to these rights in the North Takoman case will, however, worsen the democratic makeup at the overall level.

Thus far these arguments have all focused on the intrinsic worth of democracy. The wrongdoings of DR can be explained in instrumental terms as well.[83] The ability of democratic structures to mediate between competing visions and provide a responsible and legitimate way of solving disputes (Waldron, 1999: 16) is compromised by DR, and any given use of DR adds to and is indeed constitutive of this weakness. Additionally, DR institutions often take the form of allowing political actors "to be judge and jury in their own case", often insulated from the usual mechanisms of answerability and accountability so integral to democratic politics (Sutherland, 1995: 506).[84] This can be an especially severe problem if we view democracy as being justified as a way to minimise abuses and cruelty (for example, the liberalism of fear found in Shklar, 1984): the institution as a whole is unable to properly constrain such abuses, but there is still plausibly dirt that can be found in discrete cases. Acts that restrict accountability, such as gerrymandering and dishonest campaign finance systems, can erode

---

[83] Instrumental theorists of democracy can, I contend, plausibly feel dirt that comes from getting in the way of this effective instrumental operation. This is in the same way that Meisels (2008: 156-159) argues that rule-utilitarians can also feel DH problems. Besides, instrumental theorists of democracy can still sign up to there being intrinsic dirt involved with certain actions. They need not be instrumentalists about all values, after all.

[84] Sutherland uses this criticism to make the point that DH politics asks politicians to be elevated to such a position where they make life-wrecking trade-offs but are then free from the usual constraints of democratic accountability, thus a broader incompatibility between DH and democracy. The point I am making here is that sometimes those democratic institutions, owing to their flawed design and the corrupting ways they are being used, can themselves be the source of the problem, and thus become obstacles to that democratic accountability.

these protective barriers. Other dirtying acts can violate these democratic norms too, of course, but DR cases involve these violations being institutionalised. DR institutions can also entail a failure to resolve conflicts fairly (Hampshire, 1999: 18). Perhaps such institutions once solved conflicts with a reasonable degree of fairness (Hampshire, 1999: 35), but now they do not. These are all first and foremost critiques of the institutions, to be sure, but the instrumental damage here is compounded by and in large part constituted by the individual actors, even those who are ATC-justified in their participation.

DR institutions and practices also cause democratic harm by cultivating virtues and dispositions that are unsuitable for democratic politics. Ramsay (2000) argues that democratic politics are a different beast than the vicious political structures that many DH theorists assume. DH theorists, the charge goes, borrow too much from Machiavelli, and whatever the merits of what he said at the time, it is an inappropriate lens through which to discuss democracy. While it is possible to overstate this point, as quite clearly democratic politics have not eliminated either value conflicts or corrupting institutions, certain virtues and skills that were apt for Machiavelli's era are dangerous for democracies. Furthermore, the sources and consequences of political defeat were much more grave back then. A British or American leader is unlikely to be exiled by his political rivals or lose power after defeat in battle, for instance, and will have to govern with certain constitutional constraints that the Machiavellian prince would not have recognised. What DR conflicts can involve, as with DH conflicts, is the development of talents and instincts that are damaging to proper democratic functioning. DR institutions can keep alive, even in relatively minor forms, dispositions and practices that a properly functioning democratic politics should ask us to eschew. The DC of politics is made worse than it needs to be.

The existence *and* use of DR institutions thus involve violations of democratic rights and duties, cultivate undemocratic virtues, and damage the instrumental benefits of democracy. The need to avoid these costs are all addressed by pointing to a broader duty, whether grounded in natural justice or the maintenance of joint democratic projects, to abide by democratically acceptable rules. It is a duty to behave democratically. Abiding by DR institutions violates this duty. Politicians have individual responsibilities to make the joint project of democracy work. This involves removing systematic injustices and corrupting institutions, and acting in ways that support this joint project even if the institutions themselves undercut them. A duty to refrain from using DR practices thus emerges. Perfect compliance with such a duty by the collective, after all, would neuter the harm and moral dirt that comes from exploiting these institutional openings: a system that allows gerrymandering but where it is not used, for example, is flawed institutionally, but not a DR institution.

From an action-guiding standpoint, abstaining from such DR practices should be the default position, but it is a duty whose action-guiding contents can be waived particularly if the absence of collective action renders obeying it counterproductive. However, when breaking from this action-guiding content, the agent does temporarily become a perpetrator of the unjust DR system, dishonours democratic values, and in their own way makes collective duties impossible to fulfil. Their resulting democratic victories are tainted, with their intrinsic and instrumental value as a *democratic* outcome reduced.

The existence of a prior duty to make democracy work can itself be a reason to engage with DR practices. If politics is corrupt, good people need to step in and right the ship, even if this requires engaging with DR institutions. A moral imperative for reform needs reformers. A

duty to behave democratically, then, can both ground the reasons why agents should be reluctant to engage in DR acts and why, ultimately, they may need to tolerate their use.

2.2 The Emergence of Individual Duties

The duty to abide by cleaner rules is ultimately a negative duty. To meet this duty, agents should not use tools in the political toolkit that threaten the joint commitment to make democracy work, nor should they exploit opportunities created by rulesets that go against the democratic ideal. This is in addition to the positive duty to reform politics to prohibit or prevent DR tools from being used. This is where the paradox of DR arises. There is a duty to abstain from the morally harmful acts legitimised by the system, yet the system, as discussed, seems to require these acts. Duties of political success and the rights of political participation do not go away; this wrongfulness must therefore be addressed.

This wrongfulness cannot be placed solely on the structure itself, nor is it completely waived by the pull of political necessity. The actions of the agents generate an uncancelled moral disagreeableness, a *pro tanto* residue of the kind described by DH theorists. These agents cannot entirely shift the burden of responsibility onto the collective or other actors. Individual duties here have moral sway, even if of a limited kind, independently of whether others are living up to them.

Thompson's (1980) arguments for personal responsibility in "many hands" cases resonate here. DR cases defy unambiguous allocation of` causal responsibility, yet Thompson provides compelling arguments as to why we should still believe in individual responsibility. Firstly, we cannot exclusively blame those at the top of the institutional hierarchy. This ignores how those lower down the hierarchy can have considerable discretion and bargaining power, and insulates their decisions from moral accountability (Thompson, 1980: 906-907). Edna and

the North Takoma gerrymanderers need to be accountable for their own decisions, including the proportionality and intention of such acts. We also cannot say that, as it is a collective failure, we all automatically share equal responsibility (Thompson, 1980: 907-908), for this ignores the differentiating impacts of personal contributions. Edna is less responsible than a more brazen or enthusiastic agent, yet more responsible than someone not actually taking those decisions.

These "many hands" cases *can* be viewed as failures of collective action (Thompson, 1980: 908). A collective duty does not necessarily translate into individual duties unless the sufficient conditions of compliance are already in place. There might be a duty to create the conditions that trigger the necessary conditions for compliance, but that duty need not entail acting as if cleaner rules were in place. If this was the whole story, the individual dirt in DR conflicts would come from the violation of less specific duties to prevent discrete harms and to create the circumstances of compliance that trigger the emergence of individual duties.

However, I do argue that, on top of this, there is a more specific individual duty to minimise exposure to DR institutions. The existence of this duty is not necessary to show that there is a space for DR scenarios, but it is by itself sufficient, complementing these other duties to avoid discrete harms and to create the conditions for compliance. There are, in summary, three ways that individual duties, for non-decisive actors, arise in DR scenarios: the discrete wrongdoings, contributions to the triggering of collective duties, and an individual as well as a collective duty to behave democratically.

2.2.1 The Inadequacy of Collective Duties

Part of the reason that this individual duty manifests is that an exclusively collective action account does not capture all or even most DR cases. Political activity is a continuous project,

which political agents seek to shape and mould. As such, reasons for denying individual duties do not hold. To explore the distinctions between collective and individual duties, Feinberg (1970: 248) uses the example of an armed train robbery. With a well-coordinated response, a handful of the passengers could have easily and safely thwarted the robbers. As a collective entity, then, the passengers of the carriage had a duty to resist, or at least bear some kind of moral and causal responsibility for failing to do so. Absent a structure that could have coordinated this, the collective responsibility creates no individual responsibilities.

In the train robbery example, the fault is entirely with the structure—and hypothetically, we can add, any passenger who contributes to the harms of the robbery or frustrates the formation of an effective coordination structure. As Thompson (1980: 908) points out, this is emphatically not the case with political agents who do not find themselves part of an "isolated incident", such as a train robbery, but as part of an "ongoing institution" where they often possess some culpability for these wrongs and some duty to address these flaws. Participation in DR institutions is, in other words, not an episodic instance and cannot be understood as such. Provided this responsibility (however slight or indecisive) for the failure to remedy collective action problems exists, individual duties do not go away.[85]

It is also important to note that DR conflicts are not (solely) about complicity. DR *institutions* can of course create conflicts of complicity, where political actors are asked to become complicit in the wrongdoings of others who use corrupt rulesets. But the "protagonists" in DR conflicts are engaging with these practices directly themselves, using them to counter other agents in the pursuit of political projects rather than complying with these wrongful agents and rulesets to reduce harm. That is what is dirtying here. The language of complicity thus understates the involvement in harms that is created by DR conflicts.

---

[85] Such a description also applies to Structural Dirty Hands problems, a related yet distinct class of collective action problems, which I will discuss later in this chapter.

In short, there is a duty to abide by cleaner rules. This duty can be contingently grounded in the specific discrete harms that come from not doing so and by not living up to one's duties to create the conditions for triggering collective duties. It is more intrinsically grounded, however, in the fact that politics is an ongoing institution, one that political agents seek to be involved as co-authors of, and that these institutions require political practitioners to engage with these practices not to counter greater wrongs but to achieve political goods.

2.2.2 What this Duty Means

What are the practical implications of the existence of the duty to behave by cleaner rules? Firstly, this duty explains why there is dirt involved in engaging with DR *qua* DR. Agents fail to satisfy this duty whenever they engage with the dirtying practices legitimised by the DR institutions: this is the case both if this engagement is unjustified and if it is, as is the case in DR scenarios, ATC-justified. If an institution legitimises or incentivises wrongful behaviour, as is the case in the Edna and North Takoma examples, then it becomes very hard to pursue important political ends without at least in some way violating this duty.

As well as grounding a *prima facie* action-guiding presumption against using these DR behaviours, the existence of this duty also conditions how agents should approach their engagement with them when said engagement is ATC-justified. To use DR structures might well be the cost of admission into politics, but DR structures still require certain norms of behaviour, such as proportionality and just cause, that do not attach to ordinary political techniques and tactics. Edna should not approach these deals with an all-conquering zeal or a desire to win every minor battle using whatever means are available; the North Takoma legislators should limit the scope of their gerrymanders to what is necessary to restore electoral

competition. DR structures should not be treated as standard parts of the political toolkit, even if they are necessary steps for any kind of political success within a given context.

Agents such as Edna and the North Takoma legislators should acknowledge that this dirt is contingent and that the skills they are developing are not ones that should have to be cultivated within a functioning democratic polity. They should resist fully normalising or internalising the standards of politics that these institutions point to. Keeping these attitudes at the forefront of their mind and practising the self-restraint to never be entirely comfortable with their imperfect system helps guard against a common problem with reformers: that once you have won using a corrupt system, the incentive to replace it disappears.

It is helpful to view this duty to behave by cleaner rules as an imperfect duty (O'Neill, 1990; Hill, 1992; Barry and Valentini, 2009: 499-500). Perfect duties are fully specified: what they entail and to whom the duty is owed are clear from the duty itself. A duty not to introduce gerrymandering is a perfect duty in this sense. By contrast, duties about how to behave in instances where dirtying institutions already exist (as is the case for Edna and the North Takoman legislators) can be interpreted with much more ambiguity. Because the conditions in place for the perfect variants of these duties do not yet exist, these duties instead "would seem to require that agents act to bring about institutional reforms, policy changes, and indeed the creation of new agents capable of promoting the values they affirm" (Barry and Valentini, 2009: 500). The duty not to perform certain acts remains a perfect duty, but that duty can be unattainable unless the imperfect duties of reform are seen through to completion. A practice can be condemned as unjust, with the acts that it produces being seen as dirtying, but more concrete information must be provided before this leads to perfect duties.

3. Distinguishing Dirty Rules from Other Categories

To distinguish DR from other phenomena, I first propose a two-step rough-and-ready test. The first question is whether there is a discrete wrong involved in engaging with the structure. This applies even if that wrong is something relatively minor, such as entrenching an already-extant practice or cultivating dispositions and habits that would be damaging in a more just institutional framework. An affirmative answer to this question is *necessary* for DR dilemmas to exist. Agents under a DR scenario face poor moral luck of a "systemic" (Tessman, 2005) or "unnatural" (Card, 1996) kind. They become implicated in structural injustices that conflict with their own values or sense of integrity (Payson, 2017: 351), serving as "relevant, even if unintentional" contributors to these harms. The agents possess some responsibility for this structure, in the sense that it is a continuing institution that they do not wish to extricate from and whose direction they wish to shape in the service of permissible political goals.

No matter how flawed or unjust a structure is, if there is no wrongful activity in engaging with it, then no DR conflict arises. If there is nothing morally wrong about engaging with these structures, then the question is exclusively one of institutional design. The agent is spared the bad moral luck that comes from having a ruleset that is itself dirtying to play by. There can be a moral duty upon agents to rectify such structures, of course. DH dilemmas of other varieties might well arise if a political actor is motivated by moral considerations not to fix these institutions or to make them worse, but this is not a DR conflict. DR dilemmas are fundamentally about engaging *with* these structures. There must be a positive moral duty to behave as if cleaner rules were in place.

The first question, then, must be answered in the affirmative in order to create a DR conflict. The second question in this rough-and-ready test asks whether these moral costs are acceptable—for the individual *and* for the broader polity. From the individual standpoint, the

engagement with these dirtying rules must be ATC-permissible. This is not a requirement that every discrete engagement with these institutions counts as a 'lesser evil'. That is a higher standard of moral justification suitable for DE cases, but not for DR. Instead, the institutions must not be *so* warped that political engagement through these systems is no longer permissible. The assumption is that this engagement is, with due constraints such as proportionality, permissible. This engagement, after all, is the price for effective political participation.

Regarding acceptability of moral costs for the broader polity, here we distinguish DR from those cases that are better captured under the DC claim. Are these moral costs an acceptable, necessary, or even desirable element of a broader system? Less demandingly, is there a moral duty to change and resist these rules? Practices better explained by DC are those that represent an intrinsic part of the structure of political life (such as a pluralistic society) or are rooted in a contingent structure that is all-things-considered morally desirable (such as democratic politics). At the very least, DC practices are subject to reasonable disagreement. DR costs represent flawed institutions that inflict trade-offs and value costs that we should not tolerate. Those that we should tolerate should be viewed as DC instead.

For instance, many of the implications of democratic politics, such as the compromises that it requires, are to a great degree structural. To seek to "remedy" this, though, is to miss the point if we believe in democracy as having some overriding value, especially if it is of an intrinsic kind. The costs that come from democracy are thus best viewed as DC. The costs that might arise from specific incarnations of democracy and from the specific rules that govern the democratic process within a given polity are, however, potentially suitable for DR analysis. A good test for determining whether DR might arise when it comes to otherwise-worthy institutions is to ask whether the costs are necessary or proportionate for achieving the

institution's purported benefits, and whether it is morally acceptable to leave the institutions as is. Positive answers to these questions place us in the terrain of DC as opposed to DR.

The ATC-acceptability of the costs of an institution can change over time. Similarly, we might discover that a previously held answer about the acceptability of such institutions was incorrect.[86] New ideas of substantive justice can form, new answers to difficult questions can arise, and important background conditions, such as the behaviour and convictions of other actors, can change. Whether an institution's costs are justified is a question that defies final answers and should be approached with a degree of humility, both in terms of epistemic and valuational judgements.

3.1 Distinguishing Dirty Rules from Dirty Careers

In this section I will further expand upon how we can determine whether a self-justified norm of politics or some other institutional constraint falls under the purview of DR or DC. The simple answer to this question is that DR represents something pathological and remediable, and there are compelling moral reasons to change the system in question. These attributes are absent with DC. Both DR and DC represent "the rules of the game", with important moral goods and ends secured through the game. The costs of DC are accepted, though not uncritically, as part of this journey. The costs of DR should not be accepted, however. One of the moral goals in politics is to resist such institutions and change these particular rules.

Another way of presenting this is to argue that DC represents the 'core' aspects of a political role morality, which cannot or need not be changed to deliver a functioning

---

[86] Recall the discussion about "honest mistakes" and institutional mismatching from Section 1.2.1 of this chapter.

democracy.[87] A good politician will have mastery over these core dimensions of their political role morality. Peripheral aspects, in contrast, are subject to change and may (and often will) be subject to constant flux and contestation. DR dilemmas arise when those peripheral aspects become dirtying. Somebody entering politics must acknowledge that these are, *currently*, the self-justified rules of the game, but that there are valid reasons to try to change them or to at least resist fully internalising them.

3.1.1 Ambiguous Cases

Ambiguities can easily emerge. Let us add some detail to Paul's dilemma. Paul is, let us suppose, a believer in proportional representation (PR). However, his country has a first-past-the-post (FPTP) political system. Paul deeply laments this. He is realistic about the poor prospects of institutional change: even with his best efforts, it is unlikely that PR will be implemented during his political career. Paul understands that the FPTP system demands different political qualities to those that the PR system he favours would demand.

Paul views some of these qualities, in fact, as representing part of the costs of a FPTP system. For example, the FPTP system means that political success requires catering to a certain demographic who, by chance, manage to hold a decisive role in a key marginal constituency. This affects the policies and campaigning stance that Paul's party takes. Paul would rather build a broader coalition, or perhaps focus on those voters who share his values or have the most pressing moral claims. He cannot do this. He feels constrained by what he sees as the dirtying demands of the electoral system.

---

[87] I do not deny the possibility of the DR and DC distinction applying to non-democratic contexts, but I focus on democracies here.

Paul feels a real uncancelled moral cost from engaging in this prioritisation. He feels that he has let down those who he is in politics to represent and has adopted a cynical approach to policy formation that deviates from the demands of justice. Paul thus feels dirt. What kind of dirt is this? The delineable and discrete nature of DE is inappropriate here. Paul's situation is much more quotidian. To some degree, it seems like a classic case of DC. This is how we have used Paul thus far. Political success requires chasing the votes, so to speak, and while this can sometimes represent an important democratic virtue, it can also represent a moral cost.

Paul, however, does not view his predicament as being an inevitable part of politics. He has accepted that electoral success can require chasing the votes. The existence of a FPTP system, however, amplifies this cost. He now has to cater to a specific demographic not because they get him closer to winning a numerical majority, but because sheer geographic luck has given them an arbitrarily powerful say in who runs the country. The system is wrong, Paul contends, and so the dirtying demands that it asks of him are best viewed as DR.

Is he correct in this analysis? Paul is right in saying that we can imagine a political climate without the costs of FPTP. It is, institutionally speaking, an attainable goal, one that is a worthwhile political project. However, PR can make its own dirtying demands on politicians. PR systems are more likely to lead to coalitions, after all. This means compromise. Compromise, much like campaigning to specific demographics, need not be intrinsically dirtying. However, a willingness to compromise *can* be dirtying and can entail certain attitudes towards one's commitments and political supporters. Both systems, then, can make dirtying demands on politicians. These are different and perhaps even incommensurable dirtying demands, to be sure, but they are dirtying all the same. A degree of dirt is thus inevitable. Paul makes an ATC-considered judgement that he would rather incur the dirt from PR's mandated virtues and dispositions. It is not *necessarily* a reduction in dirt that he agitates for in the end, but instead a substitution of one kind for another.

There is still dirt, however. So, should the dirt from PR be viewed as DC or DR? The answer is complicated, and so the distinction between DR and DC should be viewed as possessing a fuzzy boundary, with some cases possessing characteristics of both. This case is a borderline one, where both DR and DC thinking can be applied. That there will be dirt from a chosen electoral system is a DC notion. However, the choice of one electoral system over another means that the dirt is a contingency and that it can be changed. There are therefore valid (and perhaps ATC-compelling) reasons to desire a change. For an agent who believes in that change, the mindset of DR should thus play a role. Similarly, the change must be realistic, in the sense that it is a project that can be realistically placed on the political agenda and achieved via existing political means. Whether a change is realistic in this way is more scalar.

One consequence of this analysis is that many aspects of DC can be viewed as choices, which are subject to valid political contestation, and this can entail uncomfortable losses of integrity for those who are forced to abide by rules that they disdain. What would move this instance firmly into the territory of DC is if there was no discernible right answer to be found, such that there is no moral dirt whatsoever to be found in maintaining one arrangement over another. What would move this instance firmly into DR is if the FPTP system were uniquely corrupting or of some unnecessarily onerous moral cost, and if there was a special duty to behave more akin to how a PR system would demand. Identifying the appropriate theoretical framework thus requires contextual sensitivity.

3.1.2 Can and Should Institutions Be Changed?

One interesting possibility is this: say that we agree with Paul that PR is, clearly, ATC the best electoral system. However, say that on balance its dirtying requirements for individual agents, namely its demands for compromise, inflict more dirt than those demanded by FPTP. The best

institutional arrangement is thus more dirtying. This can also be viewed as being a DC case where DR analysis can have an important role. It is a DC case in that there is no great moral reason to want to change such systems, but the DR mindset is appropriate in the sense that the dirt being accrued is a choice and an agent can permissibly seek to reform it. An agent such as Paul would be hypocritical in treating the acts required by a system that they oppose as if they were merely default costs of the political lifestyle. Sometimes it is right for politicians to treat DC cases as DR.

Can the opposite apply, and can agents treat DR cases as DC? I think so. Let us take Edna's example. What if she views the existence of the corrupt campaign finance regime as DR, as I have argued, but *de facto* fixed? She concedes that the regime is not a necessary feature of politics. She will continue advocating for change, knowing that implementing reform *is* a possibility. In effect, though, it is fixed. There is no real hope of it being changed for the duration of her political career. She must embrace this system and the dispositions to which it gives rise. To be circumspect or restrained about embracing these dispositions is to cripple her own chances for success. This cuts against how I have previously argued that agents should approach DR.

Edna would be ATC-justified in, for the most part, treating this case as she would a standard DC. This treatment of a DR institution as DC is a dirtying one, however. The threshold of justification is higher. The costs of entering a DC are higher. This should also be accompanied with an openness to moral imagination and change. Edna must be prepared to be wrong about this, and she is opening herself up to the possibility of incurring bad moral luck because of assuming that politics is even grimier than it actually is. It is also worth noting that whether an institutional arrangement is fixed is likely to, itself, be a scalar question, and settling such a question requires delving into difficult assessments of political realities, institutional design, and so on.

The *costs* of changing institutions also matter. Let us take a step back and return to Paul's example as set out in the introductory chapter. The dispositions he is being asked to cultivate, such as a certain mendacity and hypocrisy, are the result of human nature, he is told. Paul believes that human nature can be changed, however. It is possible, he believes, to have a society in which politicians can address the public honestly and with integrity. If Paul is correct, it would suggest that the demands his society places on politicians should be viewed as DR rather than DC. Paul might be right that human nature is not fixed. However, it will likely be undesirable to try and change human nature or implement a top-down restructuring of society, especially in the kind of timeframe relevant for Paul's career. The high moral costs of such actions will outweigh any moral benefits that can be delivered by potentially cutting down on future dishonesty.[88] Paul's case, then, should be viewed as DC rather than DR. Another example might be if an institutional arrangement represents a delicate compromise between plural values: navigating such an institution will create DR dilemmas in the sense that it is a contingent source of dirt and a contingent strain on the agent's integrity but will be DC if the agent judges that preserving the stability that that compromise provides is ultimately what is important. "Felt necessary evils", to reuse Levinson's terms, can thus lead to DC as well as DR.

---

[88] Barry and Valentini (2009: 508, drawing on Buchanan, 2004: 61 and Räikkä, 1998: 35) describe cases such as these as "morally inaccessible", where the moral costs needed to satisfy an ideal are "far too high". There is much that can be done via institutional reforms to mitigate the impacts of character and to remove the potential for cruelty and corruption. However, many such steps are beyond what is feasible (Shklar, 1984: 244) and the language of moral inaccessibility helps here. Though I do not think moral inaccessibility is a necessary criterion for moving something from DR to DC, it is a sufficient condition. What is important to note about this argument from infeasibility is that, while feasibility is rendered as very important for evaluating what needs to be done and how political actors should feel about their conduct in suboptimal situations, infeasibility in and of itself does not dispel dirt. Under my account, a moral rule can be morally inaccessible yet its violation still inflicts dirt. I believe this is justifiable under the structure of choice supported in Chapter 1, namely double-counted impossible oughts.

4. Structural Dirty Hands

Most discussions of political dirt have focused on two contrasting ways of understanding the source behind the messiness of politics. The DC view focuses on how the choice of pursuing political success means that one must accept rules of behaviour incompatible with ordinary morality. By contrast, the orthodox DH view (in this thesis restructured as DE) focuses on the dirt that arises from momentary paradoxes of action, often those caused by the external immoralities of others. DR, of course, is situated in a middle-ground between these extremes, where the contingent *and* remediable structural aspects of a democratic system incentivise or mandate immoral and dirtying behaviour.

Of course, DR is not the only possible way to fill in the gap between these two extremes. It is scarcely unusual for DH theorists to *acknowledge* the potential for more structural and quotidian cases of dirt that are nonetheless remediable. Walzer himself acknowledged this when he discussed the 'terms of the competition', though he seemed to walk away from this stance in his later writings. Few theorists, ultimately, *deny* the potential for structural DH. The examples they use in elucidating their concepts, however, are anything but structural, and thus certain aspects of more structural DH predicaments are not addressed within their analysis. This is why I have proposed relegating the orthodox model to merely being *part* of the DH framework, understood as DE, and allowing DR to tackle most of the structural conditions.

4.1 What are Structural Dirty Hands?

Fausto Corvino (2015) provides an alternative model of filling the gap between the two extremes. Corvino's model focuses on paradoxes of action, albeit in a more structural way. The

standard DH cases, exemplified by Ticking Bomb Scenarios, are defined by an agent being restricted by "external contingencies" beyond their control and in possession of a unique responsibility to mitigate the evils from this external immorality (Corvino, 2015: 7). Corvino draws attention to Structural DH (hereafter SDH) scenarios. In these cases, a collective action problem creates the paradox of action.

The best examples of this, Corvino argues, are the "exploit or fail" problems faced by those operating within industries that rely on sweatshop exploitation (Mayer, 2007). Suppose an executive in this industry knows that, given the current state of competition in their industry, moving away from sweatshop labour would cause his company to fail, to the benefit of his unscrupulous competitors. It would also cause the destitution, arguably a greater evil than exploitation, of the workers within the sweatshops. Unilaterally moving away from exploitative sweatshop practices, the executive concludes, is ultimately not desirable; extrication from the industry also creates greater harms. Instead, he invests considerable time and energy trying to create the conditions where moving away from exploitation becomes possible (Corvino, 2015: 9-10): lobbying government for change, trying to encourage changes in public opinion, and encouraging his competitors to agree to end exploitative labour practices. In Thompson's (1980: 908) language, the executive is trying to trigger the conditions that make individual duties kick in for all.

The executive fails in these efforts, however, and decides to continue the exploitative labour policies, as letting his company fail will cause the greatest suffering (Mayer, 2007: 616). The executive argues that the exploitation caused by his action is structural rather than "discretionary". It is not a result of his moral preferences, but instead an "inevitable result of a structural problem that he had not managed to solve together with other agents" (Corvino, 2015: 11). In conventional DH cases, agents are trapped within a "harness of necessity" (De Wijze, 2005: 467), where the best actions that can be taken are, due to external constraints, *pro*

tanto wrong. The executive, however, is not trapped in a harness of necessity of the kind De Wijze describes, in that he has not been forced into this undesirable option-set by the actions of agents that he cannot control. Instead, it is one that he has failed to disentangle himself from; a collective action problem that he has failed to solve (Corvino, 2015: 10).

DE events such as TBS scenarios represent "win-lose" scenarios (Corvino, 2015: 8), as in these cases the dilemmas cannot be avoided when every agent follows their goals and self-interest. The terrorist wants the bombs to go off and suffering to ensue, thus there is no way to get out of this predicament without either letting the terrorist achieve their evil goals or using the evil means needed to stop them. There is no collective action failure, and no way to defuse the dilemma via collective action. By contrast, the executive is facing a potential 'win-win' scenario. If all companies agree to end exploitative practices, then the competitive advantage derived from them disappears, and the "exploit or fail" DH dilemma disappears.

Instances of SDH can be viewed as arising when the regulation and normative constraints binding competition fail, resulting in an arms race of sorts in which the destructive costs of competition are unleashed (Bagg and Tranvik, 2019: 976). While the use of these unsavoury practices *originally* arose from the untempered pursuit of adversarial ethics, they are no longer *necessarily* manifestations of overzealous competition. To reverse an instance of SDH, these competitive instincts need to be reversed rather than merely regulated, and parties need to coordinate to render such practices no longer necessary for continued competitive success.

4.2 Comparing Dirty Rules to Structural Dirty Hands

Corvino's concept of SDH is an important one. Does it negate the need for DR analysis? Does it provide a sufficient answer to questions such as those facing Edna and the North Takoman

legislators? I believe the answer to both questions is 'no'. SDH is a related concept, and in part occupies the overlap between DE and DR, but DR problems such as Edna's and North Takoma's cannot be reduced to collective action problems of the sort Corvino describes.

In short, SDH cases arise when agents fail to defuse DE predicaments via coordinating a collective action solution. The justification needed to engage in a SDH practice is, like with DE, the identification of a lesser wrong. This is opposed to DR, where the justifications for abiding by contingently dirtying rulesets come in the form of ordinary political goals. Still, there can be considerable overlap. There are other distinctions too, but this is the crucial one.

4.2.1 Dirty Rules and Collective Action Solutions

Firstly, SDH cases are, akin to DR, based on bad systemic moral luck. There is a structural flaw that can be remedied. The solution to SDH dilemmas is some kind of collective action agreement, one possible within the current rules of the game, that will dissolve this specific paradox of action. In contrast, DR cases are often more entrenched, constituting the very rules of the game that themselves need remedying. To navigate away from DR institutions requires closing institutional openings, instituting wholesale changes in behaviour and practice, and preventing not only the motive but the opportunity from arising. It may not merely be difficult, but entirely impossible, to prevent DR cases with collective action solutions alone. In paradigmatic DR cases, the terms that others have set is the main source of the dilemma. In SDH cases, the behaviour of others, namely in refusing to come together to properly coordinate a solution out of the paradox of action, is the main source of the dilemma. The "harness of necessity" that constrains the feasible options available to an agent is different, representing neither the inherent rules of the game nor extraneous immoral impositions.

Another major difference is that, in Corvino's sweatshop labour example, a 'win-win' outcome exists.[89] This is a key phenomenological difference between SDH and DE. DR cases do not require such a win-win option to be available, and do not rely on it to produce their phenomenological differences with DE. In DR cases, an actor might, on balance, derive a substantial net benefit from the DR practices. Consider a party who, on net, benefits from the normalisation of gerrymandering, or the lobbying groups who predictably win more legislative battles under a corrupt campaign finance system.

In SDH cases, competitiveness requires engaging in immoral behaviour and all actors would, at least from the perspective of relative gains in comparison to one another, lose no comparative advantage from eliminating the SDH practices. This is why a collective action solution will often be unavailable in DR cases. It is usually a 'win-lose' scenario, where one actor has every reason to resist the abolition of DR practices, unlike how Corvino describes SDH cases. Despite sharing a structural source with SDH, then, many DR cases here share a similarity with more conventional DH practices, in that there can be a clear winner from these procedures.

4.2.2 Differing Necessary Conditions

The second key difference between SDH and DR is that the necessary conditions for each category are distinct. As discussed, DR dilemmas are defined not by the severity or gravity of the wrongdoing involved, but by the source and nature of what makes said dirt ATC-justified.

---

[89] One can, admittedly, dispute this as a proper characterisation of the sweatshop labour example. The advantage gained from sweatshop labour is, in part, a relative one that is needed to stay competitive. However, it might also be an absolute one: the profit margins of all involved companies will be larger with sweatshop labour than without, and that is why the executive's competitors are so unwilling to change. Of course, this suggests that abolishing sweatshop labour represents a 'lose-lose' scenario for the businesses, and Corvino could then argue that the logic here is the same as with a 'win-win' scenario.

Here, SDH differs from DR. SDH acts, like DE acts, are justified as a lesser wrongdoing. Determinate crimes are committed (putting aside the potential for long-term corruption and distortion), and determinate moral dirt is created as a result.

These acts must be justified as lesser wrongdoings, rather than as lamentable and dirtying parts of the rulesets. In the 'exploit or fail' scenario, for instance, the dirt incurred is not the corruption that comes from engaging with an unjust or suboptimal economic structure. Similarly, the dirt does not come from the fact that values and obligations are not being fully met. Neither is the dirt grounded entirely in the variable cost of the small additional harm that the actor provides. The dirt is the *evil* of exploitation and the *violation* of individuals. SDH and DR thus differ in terms of what the dirtying side of the dilemma looks like.

But what about the other side of DH dilemmas, the causes that render this dirt ATC-justified? At first glance it might seem that SDH and DR share similar justifications for incurring this dirt: morally valuable and morally permissible projects, as opposed to the higher threshold of need mandated by conventional DE, can justify the use of dirtying structures. Continued business operation seems to fall within this category, as it is certainly not the sort of compelling agent-neutral cause that merits violating the *pro tanto* norms that DE conflicts involve breaking, even if important moral ends are attached to this kind of success.

However, a more sustained analysis reveals that, in this regard, the 'exploit or fail' dynamic shares more in common with DE and less with DR. The moral heavy lifting in the example is done not by the prospect of a company failing or by a self-rationalising set of metrics of success, but instead by the idea that the workers who are being exploited otherwise face destitution. That is the real moral cost of the 'fail' side of the equation, along with the fact that the competitors have shown themselves to be opposed to the collective action solutions needed to navigate away from the SDH conflict. The 'greater evil' that is averted, then, is best understood not as business failure, but as the resulting impoverishment and (to a lesser and

more indirect degree) the reduced prospect for reform. The self-justified business goals lose their moral permissibility in such an exploitative context. The goals of averting greater suffering, however, remain salient. If the executive's sole consideration was business success, there would be no moral justification for engaging in this act.

Seeking political success will almost always be a morally permissible goal. To leave politics exclusively to the immoral is a morally alarming prospect. Leaving profit-seeking in a certain industry to the immoral is a highly regrettable predicament to be sure but does not quite have the same repugnance. Unlike with the political situations described in our DR examples, if immorality is required to succeed in an industry, the profession's role morality can no longer be morally defended. Of course, that does not dissolve the executive's predicament. He then faces a problem of extrication. The rules of the game are now so corrupt as to render the executive's continued participation, in and of itself, morally unjustifiable. However, withdrawing from sweatshop exploitation entirely may cause even greater harm. For politicians operating in a DR context, by contrast, extrication is not their moral mandate.

In conclusion, SDH and DR are rightfully understood as distinct concepts, with SDH referring to a subset of DE that are caused by the failure of agents to collectively coordinate in a way necessary to defuse the DE dilemmas. DR cases can overlap with SDH too, in cases where collective action can play a substantial role in alleviating the institutional flaws. Though SDH and DR share much in terms of source and phenomenological implications, their necessary conditions are different (though not, of course, mutually exclusive). The key difference is that SDH cases, like DE, require a determinate justification for a determinate act of wrongdoing, and thus justify and are only justified by acts or outcomes that pass this threshold.

Ultimately, I believe that SDH represents part of a middle-ground between DE and DR, an instance in which there is an overlap. There is still a discrete and determinate wrong being

committed here, and a clear paradox of action. This paradox merits focus. It generates dirt in its own right. There is a 'win-win' answer to this predicament that need not exist for DR dilemmas. However, there is still something about the "terms of competition" being corrupted, with permissible competition now requiring acts that normally count as paradoxes of action.

5. The Phenomenology of Dirty Rules

What is the phenomenology of facing a DR dilemma? The contingent and structural nature of DR means that the phenomenological frameworks suitable for DR and DE are not wholly appropriate here. In this section, I will discuss the moral emotions that DR should engender and how broader society should best view the practitioners of DR, along with the impacts that these moral emotions should have on how these practices are used.

5.1 The Appropriate Moral Emotion

Using DR should produce *some* kind of moral emotional reaction in responsible agents. They are engaging, after all, in an act of permissible moral wrongdoing and incurring moral dirt. They are breaking a morally significant duty to rise above the corrupt state of their polities and behave as if morally acceptable rules were in place. So, what is the appropriate moral emotion? As with DH, regret and remorse are inadequate. Regret ignores the agent's own deliberate causal role in what occurred, including their (partial) co-authorship of the political climate which they (rightly) decry. Remorse neglects how the agent has reason to stand by both their actions and their reasons for acting. Much like the DC politician, they cannot afford to second-guess every aspect of a decision that was taken while playing under the rules of the game as

they have found them. Likewise, Williams's notion of agent-regret, while moving in the right direction, seems to ignore both the deliberate causal role and the reasons to stand by that causal role.

We thus seem to be back in the realm of De Wijze's tragic remorse. Yet is that appropriate here? Corvino (2015: 7-12) argues that, while tragic remorse can plausibly be defended as the appropriate moral emotion in traditional DH cases (what I call DE), SDH calls for something different. In DE cases, bound by the "harness of necessity", the dirtying action is the least-bad option available to the actor. No other option is available to them, at least not without breaking another more significant moral goal. In contrast, in SDH cases, a 'win-win' way out of the dilemma is possible. Even if the agent has put considerable effort into trying to bring about the required collective action approach to achieve this, they are still pursuing what they know to be only the second-best option. They must deal with the possibility that $n + 1$ effort (n being the effort that they choose to put into escaping the SDH scenario) might have helped bring that collective action possibility into practice.

As discussed in the previous section, the 'harness of necessity' in DR cases is different from that in DE and SDH cases. In DR cases, there is a possibility of removing the structures that make political victory so morally costly, or of $n + 1$ effort nullifying the need to engage with DR practices. An integral part of tragic remorse, that an agent stands by their action and could even take a strange degree of pride in it, does not neatly translate into either SDH or DR. Granted, it has more power in DR cases where an agent might correctly believe that they are unable to break out of the contingent immoralities required by their political system. They will be able to argue that changing these rules requires engaging with them in the first place, or that the other moral goods achieved via political participation justify not focusing on such reforms. Tragic remorse thus has a role, but it should be tempered by the possibility that, once the agent had a certain degree of structural power, $n + 1$ effort could have reformed the DR institutions.

Part of why dirt is important is that it helps to frame the discussion around ideas such as accountability, forgiveness, and the broader social reckoning for acts that break important moral norms. This argument, which was laid out in Chapter 2, applies best to cases of DE, where there is a determinate act of wrongdoing and a discrete decision that can be evaluated. In cases of DE, we are more concretely situated in agent-neutral and objective moral questions. It is less apt for DC, where there are practices that must be accommodated and the dirt, while real, does not violate the norms of society's political role morality. Punishment thus seems wrong, even for a believer in the validity of punishing DH actors, and the need for the other forms of social accountability is far less important too. DR, as usual, seems to fall into the middle ground between these two: the practices and habits it cultivates are not ones that a healthy democratic society should make room for, but nor are they ones that lay outside a society's own self-justified rules.

The justifications for punishing DH cited by De Wijze do not neatly apply in DR cases. Corvino's arguments for why these do not work for SDH also apply to DR (Corvino, 2015: 11-12). The justification from catharsis has some validity, but the catharsis required will surely vary based on the effort the agent has put into changing the responsible background conditions. The justification from evil suffered applies too, but by seeking to change the background conditions, the agent already shows sensitivity to this. The language of evil will also be far too dramatic for many instances of DR, especially when it is something as insidious as institutional or individual corruption that is at stake.

Finally, causal responsibility has perhaps the most promise, and captures how although part of the dirt from DR arises from the discrete wrongdoing entailed by the usage of DR

practices, far greater condemnation is merited for those who strive to keep DR institutions in place or who use them without restraint or proportionality. To the extent that these arguments for punishing dirtying acts hold true, punishment seems like the wrong tool, owing to the self-justified and agent-relative goods that are being exchanged here and the democratic rights being exercised. However, the other kinds of social recognition of wrongdoing that have been discussed can still be justified by these arguments.[90]

5.3 Democratic Agent-Regret

"Democratic agent-regret" is one other possible way of dealing with this sort of ethical quandary. Beerbohm (2012) argues that this is the appropriate emotion in cases where we, as democratic citizens, cannot fully remove ourselves from the wrongful acts of our democratic leaders. We should seek to reduce our "complicity footprint" (Beerbohm, 2012: 10), rejecting where possible the benefits that accrue from this collective wrongdoing and accepting a moral responsibility to assist in compensating those wronged. This rests on an idea of "shared liability" (Beerbohm, 2012: 250) for democratic wrongs, even those we as individuals would have been powerless to prevent. This can be fulfilled by adopting Payson's (2017) call for meta-integrity.[91] Under meta-integrity, agents try to mitigate and manage their engagement with unjust structures, while acknowledging that complete extrication is impossible. Traditional "first-order" integrity can be impossible in a non-ideal world, but this still leaves a space for meta-integrity. Here, the agent acknowledges their causal role in harm, consciously justifies their violations, and continues to try and challenge the prevailing injustice.

---

[90] This suggests a kind of collective responsibility, and perhaps even collective punishment, whereby the whole society is deemed at fault for maintaining DR.

[91] Meta-integrity could also be useful for SDH.

This idea of democratic agent-regret could be applied to DR. Political actors have a shared responsibility for how their democracy and political climate operates. Democratic citizens also plausibly share liability for how political actors use these practices to secure and maintain power, supposedly in their benefit and for the benefit of their society. If there are corrupt features to this democracy, neither political actors nor private citizens should benefit from those unnecessarily and they should accept responsibility to make things better for those negatively impacted. DR practices should thus be used proportionally, with a responsible agent not unnecessarily or unjustifiably exacerbating the costs and harms of the flawed institutions. Agents should also make a moral commitment to, when possible, change the DR institutions and challenge the DR norms. Again, Payson's meta-integrity serves as a useful guiding concept here: not withdrawing from such structures but managing one's involvement to make tangible one's opposition.

Beerbohm uses the idea of democratic agent-regret to chart out how democratic citizens should feel about wrongful acts committed by their elected representatives. De Wijze (2018: 140-141) argues that tragic remorse is, yet again, the more appropriate moral emotion in truly ATC-justifiable cases of dirtying behaviour, for largely the same reasons as he criticises agent-regret more generally. Does this critique apply to DR? Perhaps. It depends on if the DR conflict in question is 'symmetrical' or not. If the dirt is symmetrical, tragic remorse seems to have greater hold as a recognition of the inescapability of wrongdoing as it better captures how the agent can defend the outcome of what they did.

Of course, the decision as to whether to enter a particular domain is usually asymmetrical. Once already within that career, however, DR conflicts might be better described as symmetrical, with both options causing some kind of dirt. To enter politics and to abstain from using the most effective means of achieving valuable political goals, when those hostile to one's values have no such grievances, seems to be mistaken, and an unjustifiable decision

from the standpoint of the political goods that the agent has agreed to pursue. If she does not use the DR practices available to her, Edna is letting her side be defeated by unscrupulous actors; the North Takoman legislators, too, are missing a chance to mitigate the most grievous damages of South Takoman gerrymandering. To avoid using DR in a system where they are required is to fail to take politics seriously; it is to be even more ineffectual than the purist politician who wants to avoid dirtying their hands in DE scenarios. Thus, tragic remorse can seem like the more appealing choice for DR conflicts.

Still, I think democratic agent-regret has some validity for DR. In part, this is because De Wijze's (2018: 140-141) arguments for eschewing democratic agent-regret do not apply in full here. It is true that the wrongdoings involved under DR have been given a degree of democratic licence and, from the perspective of individual political actors, are committed for justifiable and sometimes even laudable goals. However, DR cases do not arise from some laudable democratic structure and, in terms of what DR practices are being used to achieve, do not necessarily provide goods so integral as "safety and protection". They can concern "enrichment", or the pursuit of other worthy moral goals.

While Democratic DH (or DDH) and DR both involve justified acts of wrongdoing, DDH acts, when a subset of traditional DE acts, represent a tragic extraneous threat that introduces a new moral conflict that would otherwise have been absent. The justification of lesser evils is called upon to justify DDH acts. DR, by contrast, are morally unnecessary and we can plausibly, with sufficient political consensus, create a political landscape where no conflicts necessitate immorality. In the meantime, ordinary democratic functions require their deployment. As such, though each individual DR conflict might be more befitting of tragic remorse, the existence of DR conflicts within a society is more appropriately met with democratic agent-regret. The DR politician can stand by their actions yet knows that the

institution itself does not constitute permissible wrongdoing. Both moral emotions, therefore, have a role.

6. Conclusion

To understand what is at stake in cases such as Edna's and the North Takoma gerrymandering case, we need a new category of dirt in politics. The Dirty Episodes concept is certainly useful for describing cases in which there is a delineable choice that must be made, resulting in determinate dirt. With the examples of Edna and North Takoma, however, the logic of selecting a 'lesser evil' does not capture what is incentivising their active political participation, nor does the episodic structure do appropriate service to the quotidian nature of the wrongful practices in which they are engaging. Corrupt campaign finance practices and aggressive gerrymandering have become part of the self-justified ruleset of politics and making use of them cannot merely be treated as an aberrant decision. However, the Dirty Careers framework is also inappropriate here. There is something aberrational about these cases. They cannot merely be described as part of the rules of the game, as practices that the agent receives a moral blank cheque to partake in bound only by the vocation of politics itself.

This is why we need Dirty Rules as a third category of the dirt in politics. In this chapter, I have set out when and how DR dilemmas emerge and the necessary conditions for both when DR institutions come into being and when such institutions give rise to DR dilemmas. DR institutions arise when contingent structures and institutions unnecessarily and unjustifiably permit, incentivise, or otherwise legitimise corrupt or corrupting behaviour. These institutions give rise to DR dilemmas when behaving in accordance with these institutional rules is all-things-considered justified. This will often be the case when it comes to political activity. Despite this justifiability, moral dirt falls not solely on the institutions and their authors, but

also on the users of such practices. This also applies to those who are ATC-justified in accruing such dirt.

As with DC, political success under DR conditions requires engaging in these wrongful acts. Unlike with DC, this is a remediable structural feature that should be changed. There is a pressing moral duty to change the institution that causes these features. There is also a *prima facie* action-guiding duty to behave as if cleaner rules were in place, and the violation of this duty generates moral dirt. I explored the distinction between DR and DC further, considering cases where it can be hard to draw. Overall, the distinction between DR and DC dilemmas should usually be understood as scalar rather than binary, as should the prior distinction between fixed and contingent institutions. I also established the distinctions, and similarities, with Structural Dirty Hands, which were defined instances in which a paradox of action can be resolved by forming a collective action solution. Finally, I discussed the phenomenology of DR. These dilemmas can ground both democratic agent-regret and tragic remorse, and an overall need for proportionality unlike other DC acts.

CONCLUSION

There are some moral wrongs that leave a certain kind of moral residue or taint upon an agent's character. These acts and practices are dirtying. They represent unmet moral obligations for which the agent bears (at least some) moral responsibility. The costs of these acts—the moral remainders—at least in part fall upon the agent. In this thesis I have defended the controversial claim, associated both with the Walzer-inspired Dirty Hands frameworks and Machiavelli-style role morality accounts, that within politics it can be right and prudent to accrue this kind of dirt. This is not a defence of amoralism within politics, nor does it commit us to cynically accept all moral pollution within politics. However, if politics is morally important and if there is moral weight attached to both the values that politicians promote and the outcomes that they wish to engineer, then it becomes justifiable to wade into these muddy waters and get dirty.

My defence of the idea of all-things-considered justified dirt has mainly focused on identifying and responding to weaknesses in how this idea has previously been presented. The dirt of politics is usually presented by its defenders (and indeed its critics) in one of two ways: as an episodic disruption to what is otherwise an acceptable (if not harmonious) moral baseline where a specific decision requires a dirtying act, or alternatively as something that political practitioners need to buy into wholesale, where new and dirtying dispositions and virtues must be cultivated. There is merit to both accounts, and cases that both conceptualisations capture well, but neither can adequately explain the phenomenon of justifiable dirt in politics by itself.

The argument of this thesis is that we need to move away from monolithic views of how this dirt can arise. There is already a familiar way of breaking down DH along the lines of the dirt's severity. More quotidian political imperatives merit some grime, but do not legitimise more exceptional acts of wrongdoing. These more extreme acts can, however, be rendered permissible by grave moral emergencies. I endorse this way of breaking down dirt, but it is

insufficient. I have thus proposed a tripartite model of ATC-justified dirt, based around what *causes* a dirtying act or practice to become ATC-permissible. What is the source, in other words, of the "harness of necessity" (De Wijze, 2005: 467) that causes the most justifiable action to entail a *pro tanto* wrong? The answer to this question will determine how the act should be evaluated and how the agent should view its emergence. The distinction based on severity focuses on the nature of the dirt that is ATC-justified; my distinction complements this by focusing on what causes this dirt to be ATC-justified.

These three categories are Dirty Episodes, Dirty Careers, and Dirty Rules. Dirty Episodes cover those instances where a discrete decision must be taken and this decision represents something that is aberrational and flawed about a situation: something outside of normal politics has arisen, which requires stepping outside of ordinary politics to handle. Dirty Careers encapsulates the dirt that comes from accepting and internalising the self-justified goods of politics. Both DE and DC, then, represent granting a delineated role for the Walzerian and Machiavellian views of DH. Dirty Rules represent the most novel category: DR cover those instances where a remediable rule of politics, one that should and can be changed, mandates or encourages dirtying behaviour.

Under this tripartite model, acts are distinguished based on what *kind* of burden of justification is needed to make the dirtying act permissible, as opposed to a simple proportionality analysis. This helps capture the true diversity of scenarios that have been given the DH label: they all share "the idea that correct political action must sometimes conflict with profound moral norms" (Coady, 2018). DE, DC, and DR all differ on what kind of "correct political action" is sufficient to justify the dirtying behaviour, as well as the way in which the "profound moral norms" in question are experienced.

Unpacking dirt along the tripartite model helps us avoid forcing disanalogous situations into the same conceptual framework, minimising what Coady (2018: 2) described as an

unhelpful "enthusiasm to detect the sway of dirty hands" when such a model is inappropriate. By avoiding such disanalogous groupings and by encouraging a suitable sensitivity to context, we can improve the action-guiding implications of designating an instance as DH. Agents can better understand and assess the necessity (or lack thereof) of facing such choices. With this in mind, the agent can then approach the choice before them with the appropriate degree of moral imagination as to both the options they have and the (im)mutability of the circumstances. The tripartite model achieves these goals while also maintaining the benefits that can come from describing things as instances of ATC-justified dirt.

It is not enough to say that 'politicians will have to get their hands dirty by doing morally tainting things'. There are different kinds of dirt. Moral taints and moral residues can take different ways. Not everyone who has done something morally wrong will have done the same kind of thing. Not all moral rules are created equally, and the same goes for moral virtues. The same applies to the violations and departures from such norms. There are substantial quantitative and qualitative differences between different kinds of wrongdoing, and this applies just as much to justified moral wrongdoing. Many critics of DH have, understandably, attacked DH as if it does assume one type of dirt. However, shifting away from this monolithic depiction helps defuse many of these criticisms. For example, sceptics of Democratic Dirty Hands often argue that democracies delegitimise certain forms of violent behaviour in the pursuit of political ends and that they institutionally prohibit legitimate 'lone-wolf' actions. These observations are true, but only for some kinds of dirt.

1. An Overview of These Categories

A final summary of these three categories and what they cover will be useful here.

TOPIC	CHARACTERISTICS OF DE	CHARACTERISTICS OF DR	CHARACTERISTICS OF DC
	Table I. The Characteristics of Dirty Episodes, Dirty Rules, and Dirty Careers.		
When they occur	In situations of unavoidable genuine moral conflict. Bad moral luck.	In situations of genuine moral conflict, hard to avoid within a specific context, created by a contingent and remediable structure. Bad *systemic* moral luck. The choice may not be avoidable, but the overall structure is.	At moments of vocational choice ("should I enter this career, and how deeply should I immerse myself within it?"), including when adopting the mindset needed for certain kinds of political endeavour.
What they entail	Justified violation of moral norms through a "lesser evil" rhetoric: the agent does a wrongful action to prevent a greater wrong.	Justified acts of moral wrongdoing or moral corruption emanating from the use of a morally-compromising structure.	Adopting a morally compromising set of dispositions and virtues to properly participate within the political domain. The normalisation of acts that, for ordinary citizens, would count as wrongful.
What they result in	Moral residue or remainder on the agent's hands, often through some complicity in the evils and injustices committed by others.	Moral residue or remainder on the agent's hands from engaging with or perpetuating an immoral system.	Moral taint that comes from adopting a different set of values and dispositions that are required by a specific vocation. The moral loss goes beyond more than just the options not chosen but are intrinsic costs of the path chosen.
Motivation	The primary motivations to commit a justified moral violation are strong moral considerations. The avoidance of a lesser evil, for example.	The primary motivations to commit a justified moral violation are moral considerations.	The primary motivation is to embark upon a permissible, even morally important, broader political project.
Purpose of the moral violation	The moral violation seeks to bring about the lesser evil. Usually entails a very strong	The moral violation seeks to bring about the lesser evil or to proportionally enable a	The moral violations are a permissible price to pay for pursuing permissible political

	moral condemnation of the 'opposition'.	morally valuable political project.	projects.
Source	Discrete sources. Often external immorality sources. More common and weighty in politics.	They arise because of structures that impose or incentivise unnecessary moral costs.	They are an inevitable, or all-things-considered justifiable, part of a political process.

One difference worth pointing out now is the objectivity and subjectivity of the dirt involved here (though ultimately those who are agnostic on this question or favour an objective or subjective account can agree with much of my analysis). By subjectivity, I mean that the moral character of the transgression depends in large part on how the agent views their choice (impacts on feelings of integrity are key examples here, as are costs to their personal political projects and phenomenological experiences of moral emotions). How dirtying something is, under a subjective view, depends on how the agent views their moral choice. In contrast, objective views would suggest that whether an act is dirtying or not depends on an external moral judgement of their actions. So how should we describe instances of dirt? Should we view the question of whether something is dirtying, and conversely what kinds of scenarios warrant dirtying behaviour in response, as a matter of objective or subjective morality?

This is one area in which the placement on the spectrum between DC, DE, and DR particularly matters. Because DE acts need to be justified as a discrete decision aimed at securing a less wrongful outcome, they are more objective. DR acts, because they refer to wrongful acts committed in the pursuit of permissible political activity, include both objective and subjective moral ends being pursued by more objectively dirtying means. DC is the most subjective of the categories, in that while external actors might be able to say that the agent has *personally* incurred (objective or subjective) moral costs, as a whole the system is broadly acceptable so the bulk of the dirt will be expressed in the language of subjective personal commitments.

Both subjective and objective components are important throughout this spectrum, however. Ultimately both subjective and objective moral claims can be found at every stage on this spectrum. What is right for an agent to do in a particular context can depend not only on their personal roles, skills, and dispositions, but what they personally value and the norms that they themselves subscribe to. This can be a matter of external evaluation, of course, where external agents can judge the same action by different persons differently (as can be done by recognising the social value of integrity), but there is also likely to be a degree of internal subjectivity involved. Similarly, whether an agent keeps their personal integrity or maintains faith with prior particular commitments can be something that correctly affects how others, including those who do not share their substantive commitments, assess their character.[92]

2. How These Categories Interact

There are sufficient substantial differences between the categories of DC, DR, and DE to merit different treatment. There are important qualitative differences between them, and these differences create tangible differences both in how political actors should approach the prospect of accruing such dirt and how others should view said politicians. However, I do not think these are *exclusive* categories. Readers will have noticed the use of the word 'spectrum' to describe how these three categories relate to each other. This is deliberate. I do not want to suggest that we can easily place every given DH problem within one of these three categories and then authoritatively shut out attempts to draw insights from the others.[93] Instead, I believe

---

[92] For more on the importance of character, refer to Chapter II, section 4.2.3. An act that can be incredibly dirtying for one agent can be less dirtying for another; this can also include whether they have embarked on a DC or not.

[93] Such an implication would also go against previous analyses that I have made. In Chapter V, section 2.3, I argued that part of DC can include a propensity to be prepared for DE and that how an agent should handle the second-order dilemmas after a DE with any DC in mind. In Chapter VI, section 3.1.1 I discussed ambiguous cases such as how to operate in an unfair electoral system. Another example can be when the acts required by a DR institution are so severe and so dirtying as to demand DE thinking: egregious gerrymandering falls into this area.

the relationship should be viewed more like a Venn diagram of sorts, with specific examples falling into the overlaps between these ideas.

It is also worth pointing out how fuzzy boundaries arise in many of these cases. Even if we want to assign a case to one such category, what category it falls into might well be contestable. Is a political reform program desirable or morally accessible, for example? These can be hard debates to settle, yet how we answer them can determine whether something falls into a DR or DC framework. I have not offered a perfect sorting mechanism, but instead a conceptual map that helps refine and direct the needed processes of judgement and argument. The fuzziness of these boundaries can be exacerbated by the difficulty of concisely assigning individual responsibility in political systems. The president deciding whether to cooperate with an immoral regime should ideally regard it as a decision that must be made deliberately with strong consideration for the moral costs. A humble functionary implementing that decision may not have the information to draw a judgement for themself, but instead regard their choice to be in such a position, their career in other words, as what grounds whatever dirt falls upon them. Whether that role morality is something that can itself be justified, or is instead a lamentable contingency, depends on context.

3. The Five Examples Revisited

Let us now revisit the five examples laid out at the beginning of the thesis and discuss how breaking down dirt into these three categories both helps secure each example's appropriate place within DH debates and avoids flaws that come from more monolithic ways of thinking about them.

As traditionally set out, the Ticking Bomb Scenario fits most neatly with the conventional DH framework. Dana must choose whether or not to authorise a heinous act, the torture of a terrorist, in order to avert great tragedy for those that she leads. Cases such as this have been presented as *the* paradigmatic example of DH. As a result, the action-guiding and normative implications of this case have been applied to wildly different instances, or, when this application has stretched credibility too far, dissimilar cases have been pushed out of the DH model altogether.

Under the tripartite model, however, the TBS is instead located entirely within the DE category. The dirt is exceptional, the conflict a pathological one that represents a failure of ordinary politics, and both the 'disease' and the 'cure' are acts that should not be normalised as part of a morally healthy ruleset. The TBS thus covers a very important part of DH theorising, but many instances of ATC-justified dirt that fit within ordinary political functioning fall outside of it. By displacing the TBS as the paradigmatic example of *all* DH discussion, we avoid heralding the TBS as able to explain what agents should do in cases such as Paul's or North Takoma. The TBS is distinguished by the gravity of torture, the implausibility of a society having a ready-made capability to engage in these acts, and the objective severity of the evil that is needed to legitimise such actions.

As a distinct and delineable act of *pro tanto* wrongdoing, however, the TBS does serve as a clear instance of DE. Even as we place the TBS as a paradigmatic example of DE, the emphasis on contextualisation throughout this thesis still provides warnings. While it is correct to present the TBS as an anomalous paradox of action, it must still be placed within its appropriate context. The capacity to torture cannot be summoned out of thin air, nor can it be neatly contained to these cases when it is discretely justified. Similarly, decisions such as the

accountability around these acts and the demands of disclosure must be considered in relation to more traditionally political imperatives. The demands of DC and DR, in short, impact the second-order dilemmas that surround this choice, as well as the availability of and likelihood of an agent being faced with a DE such as the TBS. This has also been accompanied by a move away from the TBS as traditionally presented, owing to the limitations of such thought experiments and the implausibility of a good agent having that preexisting capacity to torture in such a scenario. Cooperation *with* torturers and extrication from torturing arrangements are more likely instances in which a dirtying choice must be made.

3.2. From Dirty Episodes to Dirty Rules: Ned, Edna, and North Takoma

Ned's case is less vulnerable to some of the conceptual problems plaguing the TBS: taking a deal, especially one that is offered, is a less exceptional act and requires a less unique skill than engaging in torture. There is of course still the risk of desensitisation to this kind of corrupt exchange, but such desensitisation is both less alarming than a disposition to torture and less likely to be intractably embedded in Ned's surrounding institutions. Still, Ned's case contains many *structural* similarities to the TBS and thus warrants placement in the DE category. Again, greater contextual awareness must be provided. In particular, the traditional Walzerian demand for Ned to show his dirtied hands and seek purification afterwards should be jettisoned, at least in the short term. For his DE act to be effective, there must be an extra degree of hypocrisy in how he presents himself going forward. His DE act must be conducted with his DC commitments in mind (and, if any arise, DR commitments too).

Edna's case, in contrast, is a paradigmatic example of a DR scenario. A remediable structural flaw, one where reform is morally accessible, creates standards of behaviour that are dirtying for her to abide by. However, she does not need to consider her every engagement

with this system on the basis of objective moral demands. She can pursue her own political projects, provided she remains cognizant of these moral costs. She can abide by the rules of the game as they are set, though she cannot go and introduce new immoralities into the system. Ned needs to justify his ends more objectively. Edna, while she has a duty to use these harmful practices proportionally and with some degree of caution, keeping an ever-watchful eye on her own contributions to these practices, can engage in these acts to perform *permissible* deeds. Ned must justify his transgression as a 'lesser evil'. Edna can engage in these practices to pursue permissible political projects. While her participation in the DR institution is dirtying, it is the price to pay for political activity. Provided she uses the DR institutions wisely and with their remediable character *in mind*, resisting their complete normalisation, she need not justify *every* move she takes within this system on a discrete 'lesser evil' basis.

If the wrongdoing Edna is involved in reaches a certain level of severity, the analysis may change. Suppose the deals that she must make have a concrete harmful impact—perhaps a variation on the school construction contracts from Ned's example, where we know the schools built will be horrendously unsafe. If these egregiously immoral trades are institutionalised, it will of course be a DR institution, but the acts it demands from agents are so dirtying that they need lesser evil justification. As a result, to remain within that system requires a greater burden of justificatory proof than the pursuit of permissible projects. It becomes something akin to the Structural Dirty Hands scenario that Corvino describes, where agents need to disentangle themselves from these scenarios and, if there is no evil that would result from their extrication, there is a *prima facie* duty to extricate. Given the power invested in politics, it is unlikely that the recommendation even in this more extreme version of Edna's case would be to abstain from politics altogether. Duties of remedy and reparation will take on a greater role, and politicians will likely need to have a serious political conviction to morally

justify their project: minimising the exposure of children to such dangerous conditions will be one such possibility.

This analysis leads to the North Takoma example. The act of gerrymandering usually needs a lesser evil justification to be permissible. Normally, this would be seen as an especially enduring instance of DE, or a kind of SDH where there is an option for a 'win-win' mutual disarmament. However, it remains firmly within the DR camp because the 'greater evil' is of an even greater loss to political competitiveness. It involves gerrymandering to reduce some of the harm created by another's act of gerrymandering. As I have argued (Sarra, 2022), in practice this means that there must be proportionality to the gerrymandering techniques, seeking to counterbalance ill-gained advantages and nothing more. Offramps should be built in, such as mechanisms for incentivising multilateral disarmament. Likewise, gerrymandering is not justified in other contexts, such as at the state level. The need to make these exceptions is highlighted by recognising both the structural nature of the wrong of gerrymandering (and thus the implausibility of containing it to discrete and delineated instances) and its remediable nature (thus both the harm of normalisation and the viability of the alternatives).

3.3. Dirty Careers and Paul

Paul's case is paradigmatic of DC: he is asked to seriously compromise his personal integrity. His choice is effectively what kind of person, as well as what kind of politician, he wants to be. Does Paul want to compromise what makes him a morally good person to pursue basic political ends? There is a degree of moral luck as to whether his dirt will end up being a worthy price to pay: despite this, the dirt is permissibly accrued. Of course, we *could* go further in Paul's case, as we did in the discussion on voting systems in Section 3.1.1 of Chapter VI and highlight particular institutions as being flawed and needing change. That this means political

effectiveness requires hypocrisy is more of a DC concern than an exclusively DR one, and these questions of integrity and hypocrisy can be relevant even for those who do not share his moral commitments regarding these voting systems. Certain aspects of DR thinking might still be useful for Paul under such scenarios: he should be keen to minimise or mitigate his co-authorship of the most harmful elements of this electoral system and to always have his eye on a long-term goal of reform.

4. Implications for Politics

I hope to have helped bolster the DH claim against various critics by breaking down this monolithic depiction of ATC-justified dirt. Agreeing with the DH thesis, as reconstructed by the tripartite model, does not require viewing all the murky realities of politics as completely immutable, nor does it require applying an antiquated overly personalised 'lone wolf' style of politics to modern liberal democratic states. DH need not arise solely within the context of stark dramatic decisions, shocking episodes representing breaks from the ordinary harmony of politics.

Throughout this thesis, I have focused on buttressing the theoretical worth of the DH claim by showing how its core component of ATC-justified dirt can arise in a variety of different ways. But what, in practical terms, does this mean for our analyses of politics? The key takeaway is this: for the politician who believes they are faced with an instance in which they can justifiably dirty their hands, they need to ask how immutable the quandary they face truly is. Is it an institution that they can fix, an inevitability of political practice, or some abnormal deviance? This is a worthwhile endeavour, and to conclude I offer some parting thoughts on why.

To answer this question, we first must remember why it is that ATC-justified moral dirt itself matters in the first place. As discussed in Chapters 1 and 2, it is both coherent and desirable to recognise that there is moral dirt inflicted in these cases. The space for tangible moral remainders is created by plausible accounts of choice in politics. Furthermore, alternative accounts that deny the possibility of ATC-justified dirt miss out on important benefits: an ability to treat certain values and deontic constraints seriously, to properly assess the broader impacts on both individual and societal characters, and to take the importance and moral weight of both politics and political values seriously. However, merely accepting dirt does not resolve the problem. An agent might incorrectly identify a problem as intractable or a dilemma as inescapable. There are reasons to believe in ATC-justified dirt, but the monolithic conception requires either a ham-fisted and damaging overapplication of the idea or restricting its scope far beyond what would be helpful.

There are reasons against viewing a DR case (such as Edna's and North Takoma) as a DE case: this encourages thinking about such decisions as if they are one-offs. In reality, DR cases cannot be judged independently from their broader context. Such treatment also obscures how the motives for Edna and the North Takoman legislators is not a lesser evil *simpliciter*. For Edna, it is about what is necessary to permissibly pursue the goods of politics. For North Takoma's politicians, it is about fighting fire with fire: using gerrymandering to help counter one of the wrongs of gerrymandering. Similarly, however, it is crucial not to treat DR cases as if they are DC. The reforms needed to end DR and change the rules of the game so that they no longer incentivise such behaviour are morally accessible. They are legitimate goals for political activity. There are therefore reasons for resisting their normalisation and maintaining some kind of integrity in the face of them. Instead of requiring an apology and the restoration of clean hands (as suggested by DE) or the normalisation of such practices and their acceptance as the innate rules of the game (as suggested by DC), DR scenarios plausibly call instead for

the adoption of some kind of meta-integrity, and a more reasoned approach to our moral violations. The DR politician may not be able to think strategically about every discrete action, and here they have something in common with DC, but they must (whether weighing up individual decisions or their broader activities) act proportionally.

The boundaries between these categories can, of course, be fuzzy. Moreover, reasonable political actors and observers could profusely disagree about what categorisation is appropriate or, more fundamentally, if we are even looking at an instance of ATC-justified dirt in the first place. Some dilemmas will thus need to be navigated with an extra degree of nuance: it may well be beneficial to apply the more demanding norms of integrity and accountability demanded by DR to a DC case, or, conversely, to allow some actors to accept a partial normalisation of DR behaviours. Other dilemmas, too, will need to be vigorously debated. Take a campaign finance system of the sort faced by Edna. Is the problem there an institutional constraint (of the kind suggested by DR), or the discrete offers and decisions presented (in a manner suggesting DE), or the underlying motivations and intractable elements of culture that lead to such corruption in the first place (suggesting DC)? As I have presented Edna's case here, it falls within DR. However, change a few empirical notes and it falls into another category.

That politics can, rightly and properly, require behaviour that generates true moral remainders is a critical observation. Political efficacy can require serious moral compromises; progressing our abstract principles in the public sphere can require deviating from them in the private. Morality is often messy. Permissible, even obligatory, choices can carry with them an uncancelled dirtying impact. However, this dirt does not take one form. The manner in which it emerges as ATC-justified ultimately helps determine how it should be viewed.

That is the main contribution of this thesis: to break down this idea of moral dirt, for the benefit of both its adherents and its critics. To do this, I have argued that both the orthodox

(or static) and the dynamic views of DH need to be used, presented as DE and DC respectively, and that neither should exclusively crowd out the other. As well as these two views, we also need the additional category of DR. DH, as a concept, covers these three ideas: paradoxes of action, the demands of politics as a different way of life, and the manner in which institutions and cultures can make politics worse. Politics is a distinct way of life, but it can still feature aberrational paradoxes of action and dysfunctional conflicts that can be remedied by morally desirable changes.

I return to how I started this thesis. We look at politics as if it is a grubby and dishonest world. Politicians are the subject of cynicism and pessimism, and talked about as if they are not suited to perform the morally important tasks that we assign to them. Quite often we decry moral crises in politics and the political class. The Dirty Hands literature has been in part an attempt by political theorists to try and make sense of this, to provide it with some kind of rationalisation or even a justification. One of the underlying aims of this thesis, then, has been to help with this task. When we decry the behaviour of politicians or politics, we should identify what exactly is causing this lamentable behaviour before committing to either acceptance or condemnation.

BIBLIOGRAPHY

Alexander, Larry (2000), "Deontology at the Threshold", *San Diego Law Review*, 37(4): 893–912.

Alexander, Larry, and Moore, Michael (2021), "Deontological Ethics", in Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*, URL = <https://plato.stanford.edu/archives/win2021/entries/ethics-deontological/> (Accessed online, 9 June 2023).

Allen, Danielle (2004), *Talking to strangers: Anxieties of citizenship since Brown v. Board of Education,* Chicago, IL: Chicago University Press.

Annas, Julia (1993), *The morality of happiness*, New York: Oxford University Press.

Anscombe, Elizabeth (1958), "Modern Moral Philosophy", *Philosophy*, 33(124): 1-19.

Archard, David (2013), "Dirty Hands and the Complicity of the Democratic Public", *Ethical Theory and Moral Practice*, 16: 777–790.

Arendt, Hannah (1967 [2006]), "Truth and politics", in *Between past and future*, London: Penguin: 223-259.

Arneson, Richard (2018), "Deontology's Travails", in Hurd, H. (Ed.), *Moral Puzzles and Legal Perplexities: Essays on the Influence of Larry Alexander*, Cambridge: Cambridge University Press: 350-370.

Athanassoulis, Nafsika (2000), "A Response to Harman: Virtue Ethics and Character Traits", *Proceedings of the Aristotelian Society (New Series)*, 100: 215–221.

Bagg, Samuel, & Tranvik, Isak (2019), "An Adversarial Ethics for Campaigns and Elections", *Perspectives on Politics*, 17(4): 973-987.

Barry, Christian, and Valentini, Laura (2009), "Egalitarian Challenges to Global

    Egalitarianism: A Critique", *Review of International Studies,* 35(3): 485-512

Beerbohm, Eric (2012), *In Our Name: The Ethics of Democracy*, Princeton University Press.

Beerbohm, Eric (2016), "The Ethics of Electioneering", *The Journal of Political Philosophy*,

    24(4): 381-405.

Beitz, Charles (2019), "How is partisan gerrymandering Unfair?", *Philosophy & Public*

    Affairs, 46(3): 323–358.

Bellamy, Alex (2004), "Supreme Emergencies and the Protection of Non-Combatants in

    War", *International Affairs (Royal Institute of International Affairs 1944-)*, 80(5):

    829-850.

Bellamy, Alex (2007), "Dirty Hands and Lesser Evils in the War on Terror", *British Journal*

    of Politics and International Relations, 9: 509-526.

Bellamy, Richard (2010), "Dirty hands and clean gloves", *European Journal of Political*

    Theory, 9(4): 412-430.

Bellamy, Richard (2012), "Democracy, Compromise and the Representation Paradox:

    Coalition Government and Political Integrity", *Government and Opposition*, 47(3):

    441–465.

Berlin, Isaiah (1980), "The Originality of Machiavelli", in Hardy, H. (Ed.), *Against the*

    Current: Essays in the History of Ideas, New York: Viking: 20–73.

Brandt, Richard (1972), "Utilitarianism and the Rules of War", *Philosophy & Public Affairs*,

    1(2): 145-165.

Brecher, Bob (2007), *Torture and the ticking bomb*, Oxford: Wiley-Blackwell.

Brettschneider, Corey (2007), *Democratic Rights: The Substance of Self-Government*, Princeton: Princeton University Press.

Buchanan, Allen (2004), *Justice, Legitimacy, and Self-Determination*, Oxford: Oxford University Press.

Calder, Todd (2013), "Is evil just very wrong?", *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 163(1): 177-196.

Calhoun, Cheshire (1995), "Standing for Something", *The Journal of Philosophy*, 92(5): 235-260.

Camus, Albert (1949 [1958]), "The Just Assassins", in Gilbert, S. (trans.), *Caligula and Three Other Plays*, New York.

Card, Claudia (1996), *The unnatural lottery: Character and moral luck*, Philadelphia: Temple University Press.

Card, Claudia (2002), *The atrocity paradigm*, Oxford: Oxford University Press.

Childress, James (1982), *Moral Responsibility in Conflicts*, Baton Rouge: Louisiana State University Press.

Chisholm, Roderick (1963), "Supererogation and Offence: A Conceptual Scheme for Ethics", *Ratio*, 5: 1–14.

Coady, C.A.J. (1990), "Messy Morality and the Art of the Possible," *Proceedings of the Aristotelian Society*, Supp. Vol. 1990: 263.

Coady, C.A.J. (2008), *Messy Morality: the Challenge of Politics*, Oxford: Oxford University Press.

Coady, C.A.J. (2018), "The Problem of Dirty Hands", in Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*, URL = <https://plato.stanford.edu/archives/spr2014/entries/dirty-hands/> (Accessed online, 9 June 2023).

Cole, Kevin (2018), "Real-World Criminal Law and the Norm Against Punishing the Innocent: Two Cheers for Threshold Deontology", in Hurd, H. (Ed.), *Moral Puzzles and Legal Perplexities: Essays on the Influence of Larry Alexander*, Cambridge: Cambridge University Press: 388-406.

Corvino, Fausto (2015), "Punishing Atypical Dirty Hands: Assessing the Moral Value of Coordination Failure", *International Journal of Applied Philosophy*, 29(2): 281-297.

Cullity, Garrett (2007), "The Moral, the Personal, and the Political" in Primoratz, I. (Ed.), *Politics and Morality*, New York: Palgrave Macmillan: 54-75.

Curzer, Howard (2006), "Admirable Immorality, Dirty Hands, Ticking Bombs, and Torturing Innocents", *The Southern Journal of Philosophy*, 44(1): 31-56.

Dancy, Jonathan (2004), *Ethics without Principles*, Oxford: Oxford University Press.

Day, J.P. (1989), "Compromise", *Philosophy*, 64: 471–485.

DeLapp, Kevin (2009), "Les Mains Sales Versus Le Sale Monde: A Metaethical Look at Dirty Hands", *Essays in Philosophy*, 10(1):13.

De Wijze, Stephen (2002), "Defining Evil: Insights from the Problem of "Dirty Hands"", *The Monist*, 85(2): 210-238.

De Wijze, Stephen (2005), "Tragic-Remorse — The Anguish of Dirty Hands", *Ethical Theory and Moral Practice*, 7(5): 453-471.

De Wijze, Stephen (2007), "Dirty Hands: Doing Wrong to do Right", in Primoratz, I. (Ed.), *Politics and Morality*, New York: Palgrave Macmillan: 3-19.

De Wijze, Stephen (2009), "Targeted killing: a 'dirty hands' analysis", *Contemporary Politics*, 15(3): 305-320.

De Wijze, Stephen (2012), "The Challenge of a Moral Politics: Mendus and Coady on Politics, Integrity and 'Dirty Hands'", *Res Publica*, 18: 189-200.

De Wijze, Stephen (2013), "Punishing Dirty Hands: Three Justifications", *Ethical Theory and Moral Practice*, 16(4): 879-897.

De Wijze, Stephen (2018), "The Problem of Democratic Dirty Hands: Citizen Complicity, Responsibility, and Guilt", *The Monist*, 101: 129–149.

De Wijze, Stephen, and Goodwin, Tom (2009), "Bellamy on Dirty Hands and Lesser Evils: A Response", *The British Journal of Politics and International Relations*, 11: 529-540.

Donagan, Alan (1977), *The Theory of Morality*, Chicago: University of Chicago Press.

Dovi, Suzanne (2005), "Guilt and the Problem of Dirty Hands", *Constellations*, 12(1): 128-146.

Driver, Julia (1992), "The Suberogatory", *Australasian Journal of Philosophy*, 70: 286–295.

Dworkin, Ronald (2001), "Do Liberal Values Conflict?", in Dworkin, R., Lilla, M., and Silvers, R. (Eds.), *The Legacy of Isaiah Berlin*, New York: New York Review of Books.

Eberle, Christopher (2021), "Irreconcilable Disagreement: Supreme Emergency, Respect, and Restraint", *Social Theory and Practice*, 47(3): 457-484.

Edyvane, Derek (2011), "Tolerance and pain", *Critical Review of International Social and Political Philosophy*, 14(3): 405-419.

Edyvane, Derek (2019), "'Richly Imaginative Barbarism': Stuart Hampshire and the Normality of Conflict", *Theoria*, Issue 160, 66(3): 9-26.

Feinberg, Joel (1970), *Doing and Deserving*, Princeton, N.J.: Princeton University Press.

Finlay, Christopher (2011), "Dirty hands and the romance of the ticking bomb terrorist: a Humean account", *Critical Review of International Social and Political Philosophy*, 14(4): 421-442.

Formosa, Paul (2008), "A conception of evil", *The Journal of Value Inquiry*, 42: 217–239.

Forrester, Mary (1975), "Some Remarks on Obligation, Permission, and Supererogation", *Ethics*, 85: 219–226.

Forrester, Katrina (2012), "Judith Shklar, Bernard Williams and political realism", *European Journal of Political Theory*, 11(3): 247–272.

Fried, Charles (1976), "The Lawyer as Friend: The Moral Foundations of the Lawyer-Client Relation", *Yale Law Journal*, 85: 1060-1089.

Galston, William (2005), *The Practice of Liberal Pluralism*, Cambridge: Cambridge University Press.

Galston, William (2010), "Realism in political theory", *European Journal of Political Theory*, 9(4): 385–411.

Gardner, John (2004), "The Wrongdoing That Gets Results", *Philosophical Perspectives*, 18: 53-88.

Gardner, John, and Macklem, Timothy (2002), "Reasons", in Coleman, J., Himma, K. E., and Shapiro, S. (Eds.), *The Oxford Handbook of Jurisprudence and Philosophy of Law*, Oxford: Oxford University Press: 440-475.

Garrett, Stephen (1994), "Political Leadership and the Problem of "Dirty Hands"", *Ethics & International Affairs*, 8: 159-175.

Goldberg, Zachary (2019), "A Relational Approach to Evil Action: Vulnerability and its Exploitation", *The Journal of Value Inquiry*, 53: 33-53.

Goodwin, Tom (2009), *The problem of dirty hands: examining and defending a special case of inescapable moral wrongdoing*, PhD Thesis University of Manchester.

Gowans, Christopher (1994), *Innocence Lost: An Examination of Inescapable Moral Wrongdoing*, Oxford: Oxford University Press.

Griffin, Leslie (1989), "The Problem of Dirty Hands", *The Journal of Religious Ethics*, 17(1): 31-61.

Griffin, Leslie (1995), "The Lawyer's Dirty Hands", *Georgetown Journal of Legal Ethics*, 8(2).

Gutmann, Amy, and Thompson, Dennis, (2012), *The Spirit of Compromise: Why Governing Demands It and Campaigning Undermines It*, Princeton, NJ: Princeton University Press.

Hall, Edward (2022), "Political Compromise and Dirty Hands", *The Review of Politics*, 84(2): 214-237.

Hampshire, Stuart (1989), *Innocence and Experience*, Cambridge: Cambridge University Press.

Hampshire, Stuart (1996), "Justice Is Conflict: The Soul and the City", *The Tanner Lectures on Human Values*, URL = https://tannerlectures.utah.edu/_resources/documents/a-to-z/h/Hampshire98.pdf (Accessed online, 9 June 2023).

Hampshire, Stuart (1999), *Justice Is Conflict*, Princeton, NJ: Princeton University Press.

Hardin, Garrett (1974), "Commentary: Living on a Lifeboat", *BioScience*, 24(10): 561-568.

Hare, R. M (1972), "Rules of War and Moral Reasoning", *Philosophy & Public Affairs*, 1(2): 166-181.

Hill, Thomas (1992), *Dignity and Practical Reason in Kant's Moral Theory*, Ithaca: Cornell University Press.

Hollis, Martin (1982), "Dirty hands", *British Journal of Political Science*, 12(4): 385–398.

Howard, Jeffrey (2019), "The labors of justice: democracy, respect, and judicial review", *Critical Review of International Social and Political Philosophy*, 22(2): 176-199,

Hurka, Thomas (1996), "Monism, Pluralism, and Rational Regret", *Ethics*, 106(3): 555-575.

Hursthouse, Rosalind, and Pettigrove, Glen (2016), "Virtue Ethics", in Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*, URL = <https://plato.stanford.edu/archives/win2016/entries/ethics-virtue/> (Accessed online, 9 June 2023).

Issacharoff, Samuel (2002), "Gerrymandering and Political Cartels", *Harvard Law Review*, 116(2): 593-648.

Johnson, Peter (1993), *Politics, innocence and the limits of goodness*, London: Routledge.

Johnson, Christa (2019), "How Deontologists Can Be Moderate (and Why They Should Be)", *The Journal of Value Inquiry*, 54: 227–243.

Johnson, Robert and Cureton, Adam (2017), "Kant's Moral Philosophy", in Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy,* URL = <https://plato.stanford.edu/archives/fall2017/entries/kant-moral/> (Accessed 9 June 2023).

Kamm, Frances (2000), "The Doctrine of Triple Effect and Why a Rational Agent Need Not Intend the Means to His End", Part I, *Proceedings of the Aristotelian Society, Supplementary Volumes*, 74: 21–39.

Kang, Michael (2020), "Hyperpartisan Gerrymandering", *Boston College Law Review*, 61(4): 1379-1445.

Kekes, John (1990), *Facing evil*, Princeton: Princeton University Press.

Kellison, Rosemary (2019), *Expanding Responsibility for the Just War: A Feminist Critique*, Cambridge: Cambridge University Press.

Kellison, Rosemary (2021), "The Critical Power of an Expanded Concept of Moral Injury", *The Journal of Religious Ethics*, 49(3): 442-461.

Kim, Sung Ho (2022), "Max Weber", in Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*, URL = <https://plato.stanford.edu/archives/win2022/entries/weber/> (Accessed 9 June, 2023).

Klockars, Carl (1980), "The Dirty Harry Problem", *The Annals of the American Academy of Political and Social Science*, 452, The Police and Violence: 33-47.

Kramer, Matthew (2014), *Torture and Moral Integrity: A Philosophical Enquiry*, Oxford: Oxford University Press.

Kramer, Matthew (2018), "Problems of Dirty Hands as a Species of Moral Conflicts", *The Monist*, 101: 187–198.

Lawrence, Mark (1990), "Legislative Ethics: Improper Influence by a Lawmaker on an Administrative Agency," *Maine Law Review*, 42: 423-452.

Lepora, Chiara, and Goodin, Robert (2013), *On Complicity and Compromise*, Oxford: Oxford University Press.

Levin, Ronald (1996), "Congressional Ethics and Constituent Advocacy in an Age of Mistrust", *Michigan Law Review*, 95(1): 1-110.

Levinson, Sanford (2020), "Celebrating the Founders or Celebrating the Constitution: Reflections on Constitution Day, 2019", *Northeastern University Law Review*, 12(2): 375-391.

Levy, Neil (2007), "Punishing the Dirty", in Primoratz, I. (Ed.), *Politics and Morality*, New York: Palgrave Macmillan: 38–53.

Liberto, Hallie (2014), "The Moral Specification of Rights: A Restricted Account", *Law and Philosophy*, 33(2): 175-206.

Liberto, Hallie, and Harrington, Fred (2017), "Evil, wrongdoing, and concept distinctness", *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 173(6): 1591-1602.

Lovell, David (2007), "Lying and Politics", in Primoratz, I. (Ed.), *Politics and Morality*, New York: Palgrave Macmillan: 189–208.

Luban, David (1980), "The Romance of the Nation-State", *Philosophy & Public Affairs*, 9(4): 392-397.

Luban, David (1985), "Bargaining and compromise: Recent work on negotiation and informal justice", *Philosophy and Public Affairs*, 14: 397–416.

Luban, David (1988), *Lawyers and Justice: An Ethical Study*, Princeton: Princeton University Press.

Luban, David (1990), "Freedom and Constraint in Legal Ethics: Some Mid-course Corrections to Lawyers and Justice", *Maryland Law Review*, 49(2): 424-462.

Lukes, Steven (2006), "Liberal Democratic Torture", *British Journal of Political Science*, 36(1): 1-16.

Lyons, Johnny (2023), "Morality, politics, and contingency", *European Journal of Philosophy*, 31(1): 179–194.

Machiavelli, Niccolò (1532 [2008]), *The Prince*, Peter Bondanella (trans.), Oxford: Oxford University Press.

Maclntyre, Alasdair (1966), *A Short History of Ethics*, New York: Macmillan Publishing.

MacIntyre, Alasdair (2007), *After Virtue: A study in moral theory*, University of Notre Dame Press, USA.

Margalit, Avishai (2012), *Compromise and rotten compromises*, Princeton, NJ: Princeton University Press.

Mayer, Robert (2007), "Sweatshops, Exploitation, and Moral Responsibility," *Journal of Social Philosophy*, 38(4): 612–6.

McDonald, Michael (2000), "Hands: Clean and Tied or Dirty and Bloody?", in Rynard, P., and Shugarman, David. (Eds.), *Cruelty & Deception: The Controversy Over Dirty Hands in Politics*, Ontario: Broadview Press: 187-197.

McKeown, Maeve (2021), "Structural injustice", *Philosophy Compass*, 16(7).

McMahan, Jeff (2006), "Torture, morality and the law", *Case Western Reserve Journal of International Law*, 37: 241–248.

McNamara, Paul (2011), "Supererogation, Inside and Out: Toward an Adequate Scheme for Common-Sense Morality", in Timmons, M. (Ed.), *Oxford Studies in Normative Ethics (Vol. 1)*, Oxford: Oxford University Press.

Meisels, Tamar (2008), "Torture and the Problem of Dirty Hands", *Canadian Journal of Law and Jurisprudence*, 21(1): 149–173.

Mendus, Susan (1988), "The Serpent and the Dove", *Philosophy*, 63(245): 331-343.

Mendus, Susan (2009), *Politics and Morality*, Cambridge: Polity Press.

Miller, Seumas (2007), "Noble Cause Corruption in Politics", in Primoratz, I. (Ed.), *Politics and Morality*, New York: Palgrave Macmillan: 92–112.

Moore, Michael (1989), "Torture and the Balance of Evils", *Israel Law Review*, 23(2-3): 280-344.

Moore, Michael (1997), *Placing Blame: A General Theory of the Criminal Law*, Oxford: Oxford University Press.

Moore, Michael (2018), The Rationality of Threshold Deontology, in Hurd, H. (Ed.), *Moral Puzzles and Legal Perplexities: Essays on the Influence of Larry Alexander* (pp. 371-387). Cambridge: Cambridge University Press.

Morgenthau, Hans (1933 [2012]), *The Concept of the Political*, H. Behr and F. Rösch (Eds.), London: Palgrave Macmillan.

Morton, Adam (2004), *On Evil*, New York: Routledge.

Nagel, Thomas (1976), "Moral Luck", *Proceedings of the Aristotelian Society, Supplementary Volumes*, 50: 137-151.

Nagel, Thomas (1986), *The View From Nowhere*, New York: Oxford University Press.

Nagel, Thomas (1993), "Ruthlessness in public life", in *Mortal Questions*, Cambridge: Cambridge University Press: 75–90.

Nelkin, Dana (2023), "Moral Luck", in Zalta, E. and Nodelman, U. (Eds.), *The Stanford Encyclopedia of Philosophy*, URL = <https://plato.stanford.edu/archives/spr2023/entries/moral-luck/> (Accessed online 10 June 2023).

Nick, Christina (2019a), "Can our Hands Stay Clean?", *Ethical Theory and Moral Practice*, 22(4): 925-940.

Nick, Christina (2019b), "In Defence of Democratic Dirty Hands", *Theoria*, 160, 66(3): 71-94.

Nick, Christina (2022), "Official Apologies as Reparations for Dirty Hands", *Journal of Social Philosophy*, 1– 16.

Nielsen, Kai (2007), "There is no Dilemma of Dirty Hands", in Primoratz, I. (Ed.), *Politics and Morality*, New York: Palgrave Macmillan: 20-37.

Nussbaum, Martha (2000), *Women and human development: the capabilities approach*, Cambridge New York: Cambridge University Press.

Nussbaum, Martha (2011), *Creating Capabilities*, Cambridge, Mass: Belknap.

O'Neill, Onora (1990), *Constructions of Reason: Explorations of Kant's Practical Philosophy*, Cambridge: Cambridge University Press.

Osborne, Thomas (2022), "Tragedy, Modernity and Political Life", *Society*, 59: 529–536.

Payson, Jessica (2017), "The "Meta" Level of Integrity: Integrity in the Context of Structural Injustice", *Hypatia*, 32(2): 347-362.

Philp, Mark (2007), *Political Conduct*, Cambridge, Mass: Harvard University Press.

Philp, Mark (2012), "Realism without Illusions", *Political Theory*, 40(5): 629-649.

Primoratz, Igor, (ed.) (2007), *Politics and Morality*, New York: Palgrave Macmillan.

Provis, Chris (2005), "Dirty Hands and Loyalty in Organisational Politics", *Business Ethics Quarterly*, 15(2): 283-298.

Räikkä, Juha (1998), "The Feasibility Condition in Political Theory", *The Journal of Political Philosophy*, 6: 27-40.

Ramsay, Maureen (2000), "Democratic Dirty Hands", in Cliff, L., Ramsay, M., and Bartlett, D. (Eds.), *The Politics of Lying – Implications for Democracy*, Basingstoke: Macmillan: 27–42.

Ramsey, Paul (1961), *War and the Christian Conscience*, Durham: Duke University Press.

Rasmussen, Larry (1972), *Dietrich Bonhoeffer: Reality and Resistance*, Nashville: Abingdon Press.

Richards, David (1971), *A Theory of Reasons for Action*, Oxford: Clarendon Press.

Roadevin, Cristina (2019), "To Punish or to Forgive?: Responding to Dirty Hands in Politics", *Theoria*, 160, 66(3): 122-142.

Roca, Esther (2010), "The Exercise of Moral Imagination in Stigmatized Work Groups", *Journal of Business Ethics*, 96: 135–147.

Rohlf, Michael (2020), "Immanuel Kant", in Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*, URL = <https://plato.stanford.edu/archives/fall2020/entries/kant/> (Accessed online 10 June 2023).

Ross, W.D. (1930), *The Right and the Good*, Oxford: Clarendon Press.

Rubenstein, Jennifer (2015), *Between Samaritans and States: the Political Ethics of Humanitarian INGOs.*, Oxford: Oxford University Press.

Russell, Luke (2007), "Is Evil Action Qualitatively Distinct from Ordinary Wrongdoing?", *Australasian Journal of Philosophy*, 85(4): 659-677.

Sandsmark, Evan (2022), "Dirty Hands, Supreme Emergencies, and Catholic Moral Theology", *The Journal of Religious Ethics*, 49(4): 739-767.

Sarra, Gianni (2022), "Fighting fire with fire: the ethics of retaliatory gerrymandering", *Critical Review of International Social and Political Philosophy*, DOI: 10.1080/13698230.2022.2056354.

Sartre, Jean Paul (1949), "Dirty Hands", in *No Exit and Three Other Plays*, Lionel Abel (trans.)., New York: Vintage Books.

Scarry, Elaine (2004), "Five Errors in the Reasoning of Alan Dershowitz", in Levinson, S. (Ed.), *Torture - A Collection*, Oxford and New York: Oxford University Press.

Sen, Amartya (1985), *Commodities and Capabilities*, New Delhi: Oxford University Press.

Shklar, Judith (1984), *Ordinary vices*, Harvard University Press: Cambridge.

Shklar, Judith (1990), *The faces of injustice*, New Haven: Yale University Press.

Shotwell, Alexis (2016), *Against Purity: Living Ethically in Compromised Times*, University of Minnesota Press.

Shue, Henry (1978), "Torture", Philosophy and Public Affairs, 7(2): 124-143.

Shue, Henry (2006), "Torture in Dreamland: Disposing of the Ticking Bomb", *Case Western Reserve Journal of International Law*, 37(2): 231-239.

Shugarman, David (2000), "Democratic Dirty Hands?", in Rynard, P., and Shugarman, D. (Eds.), *Cruelty and Deception – The Controversy over Dirty Hands in Politics*, Ontario: Broadview Press: 229–250.

Sigwart, Hans-Jörg (2013), "The logic of legitimacy: Ethics in political realism", *The Review of Politics,* 75: 407–432.

Singer, Marcus (2004), "The concept of evil", *Philosophy*, 79: 185–214.

Skerker, Michael (2020), "What Can Be Asked of Interrogators?", in Barela, S., Fallon, M., Gaggioli, G., and Ohlin, J. D. (Eds.), *Interrogation and Torture: Integrating Efficacy with Law and Morality*, Oxford: Oxford University Press: 253-278.

Skinner, Quentin (1978), *The Foundations of Modern Political Thought, (Volume I: The Renaissance)*, Cambridge: Cambridge University Press.

Smith, Patrick Taylor (2018), "Political Revolution as Moral Risk", *The Monist*, 101: 199–215.

Sorell, Tom (2015), "The Dogma of the Priority of Private Morality", *American Philosophical Quarterly*, 52(1): 89-101.

Sreenivasan, Gopal (2002), "Errors about Errors: Virtue Theory and Trait Attribution", *Mind*, 111: 47–68.

Stocker, Michael (1990), *Plural and Conflicting Values*, Oxford: Oxford University Press.

Sutherland, S. L. (1995), "The Problem of Dirty Hands in Politics: Peace in the Vegetable

    Trade", *Canadian Journal of Political Science*, 18(3): 479–507.

Sutherland, S.L. (2000), "Retrospection and Democracy: Bringing Political Conduct Under

    the Constitution," in Rynard, P., and Shugarman, D. (eds.), *Cruelty and Deception –*

    The Controversy over Dirty Hands in Politics, Ontario: Broadview Press: 207-227.

Tessman, Lisa (2005), *Burdened Virtues*, New York: Oxford University Press.

Thalos, Mariam (2018), "Dirty Hands: The Phenomenology of Acting As an Authorized

    Agent", *The Monist*, 101: 170–186.

Thompson, Dennis (1980), "Moral responsibility of public officials: The problem of many

    hands", *American Political Science Review*, 74(4): 905-915.

Thompson, Dennis (1987), *Political Ethics and Public Office*, Cambridge Mass.: Harvard

    University Press.

Thompson, Janna (2007), "Political Complicity: Democracy and Shared Responsibility", in

    Primoratz, I. (Ed.), *Politics and Morality*, New York: Palgrave Macmillan: 153–169.

Thomson, Judith Jarvis (1976), "Self-Defense and Rights," *Lindley Lecture at the University*

    of Kansas, Lawrence: University of Kansas Press.

Tillyris, Demetris (2013), *Political Morality and the Problem of Dirty Hands: A*

    Philosophical Critique and Re-interpretation, PhD Thesis University of Leeds.

Tillyris, Demetris (2015), "'Learning How Not to Be Good': Machiavelli and the Standard

    Dirty Hands Thesis", Ethical Theory and Moral Practice, 18(1): 61-74.

Tillyris, Demetris (2016a), "After the Standard Dirty Hands Thesis: Towards a Dynamic

    Account of Dirty Hands in Politics", *Ethical Theory and Moral Practice*, 19: 161-175.

Tillyris, Demetris (2016b), "The virtue of vice: a defence of hypocrisy in democratic politics", *Contemporary Politics*, 22(1): 1-19.

Tillyris, Demetris (2017), "Political Integrity and Dirty Hands: Compromise and the Ambiguities of Betrayal", *Res Publica*, 23: 475-494.

Tillyris, Demetris (2018), "Reflections on a Crisis: Political Disenchantment, Moral Desolation, and Political Integrity", *Res Publica*, 24: 109-131.

Tillyris, Demetris (2019a), "Political Realism and Dirty Hands: Value Pluralism, Moral Conflict and Public Ethics", *Philosophia*, 47: 1579-1602.

Tillyris, Demetris (2019b), "Dirty Hands and Suffering", *Theoria*, 160, 66(3): 95-121.

Waldron, Jeremy (1999), *Law and Disagreement*, Oxford: Oxford University Press.

Waldron, Jeremy (2005), "Torture and Positive Law: Jurisprudence for the White House", *Columbia Law Review*, 105(6): 1681-1750.

Waldron, Jeremy (2011), "What Are Moral Absolutes Like?", *The Harvard Review of Philosophy*, 18(1): 4-30.

Waldron, Jeremy (2018), "Dirtying One's Hands by Sharing a Polity with Others", *The Monist*, 101: 216–234.

Walker, Margaret (2003), *Moral contexts*, Lanham, Md.: Rowman and Littlefield.

Walzer, Michael (1973), "Political Action: The Problem of Dirty Hands", *Philosophy and Public Affairs*, 2(2): 160–180.

Walzer, Michael (1977), *Just and Unjust Wars*, New York: Basic Books, Inc.

Walzer, Michael (2004), *Arguing About War*, New Haven: Yale University Press.

Walzer, Michael (2006), "Terrorism and Just War", *Philosophia*, 34: 3–12.

Weber, 1919 [1946], "Politics as a Vocation", from Gerth, H., and Mills, C. W. (Eds. and trans.), *Max Weber: Essays in Sociology*, New York: Oxford University Press: 77-128.

Werhane, Patricia (1998), "Moral Imagination and the Search for Ethical Decision-Making in Management", *Business Ethics Quarterly*, Special issue 1: 75–98.

Werhane, Patricia (1999), *Moral Imagination and Management Decision-Making*, New York: Oxford University Press.

Werhane, Patricia (2015), "Moral Imagination", *Wiley Encyclopedia of Management*, 2: 1-2.

Wertheimer, Alan (1987), *Coercion*, Princeton, N.J.: Princeton, New Jersey: Princeton University Press.

Williams, Bernard (1973a), "Ethical consistency", in *Problems of the Self: Philosophical Papers 1956–1972*, Cambridge: Cambridge University Press: 166-186.

Williams, Bernard (1973b), "A Critique of Utilitarianism", in Williams, B., and Smart, J.J.C., (eds.), *Utilitarianism: for and against*, Cambridge: Cambridge University Press: 77–150.

Williams, Bernard (1976), "Moral Luck", *Proceedings of the Aristotelian Society, Supplementary Volumes*, 50: 115-135.

Williams, Bernard (1978), "Politics and Moral Character", in Hampshire, S. (Ed.), *Public and Private Morality*, Cambridge: Cambridge University Press: 55-73.

Williams, Bernard (1994), *Shame and Necessity*, London: University of California Press.

Williams, Bernard (1995), "Professional morality and its dispositions", in *Making Sense of Humanity: And Other Philosophical Papers 1982–1993*, Cambridge: Cambridge University Press: 192-202.

Wolgast, Elizabeth (1992), *Ethics of an Artificial Person*, Stanford University Press.

Wojtowicz, Jake (2022), "The Purity of Agent-Regret", *Philosophy*, 97(1): 71-90.

Wolin, Sheldon (2004), *Politics and vision: Continuity and innovation in Western political thought*, Princeton, NJ: Princeton University Press.

Yemini, Moran (2014), "Conflictual Moralities, Ethical Torture: Revisiting the Problem of "Dirty Hands"", *Ethical Theory and Moral Practice*, 17(1): 163-180.

Yeo, Michael (2007), "On the One Hand and On the Other" in Rynard, P., and Shugarman, David. (Eds.), Cruelty & Deception: The Controversy Over Dirty Hands in Politics, Ontario: Broadview Press: 157-173.

Yoder, John (1964), *Christian Witness to the State*, Newton, KS: Faith and Life Press.

Yoder, John (1984), *The Priestly Kingdom*, Notre Dame: University of Notre Dame Press.

Yoder, John (1985), *What Would You Do?*, Kitchener, Ontario: Herald Press.

Young, Iris Marion (2011), *Responsibility for justice*, Oxford University Press.

Zamir, Eyal, and Medina, Barak (2010), *Law, Economics, and Morality*, Oxford University Press.