The Journal of Ethics (2023) 27:495–514
https://doi.org/10.1007/s10892-023-09444-7

When the Heavens Fall: The Unintelligible and the Unthinkable

Lisa Tessman[1]

Received: 24 August 2022 / Accepted: 25 July 2023 / Published online: 12 August 2023
© The Author(s), under exclusive licence to Springer Nature B.V. 2023

Abstract
Moral dilemmas are a feature of moral life that make us vulnerable to tragic failures. But while all moral dilemmas involve unavoidable moral failure and leave a moral remainder, they do not all involve dirty hands. Recognizing that Thomas Nagel's ideas about the availability of both agent-relative and agent-neutral perspectives from which to ask moral questions formed the backdrop to Michael Walzer's work on dirty hands fifty years ago, this paper tries to explain why, when we must take both perspectives—thus being an agent who meanwhile considers overriding the responsibilities of our particular agency—we risk dirtying our hands. The answer to a question about what ought to happen (as judged from an agent-neutral perspective) and the answer to a question about what one ought to do (as judged from an agent-relative perspective) are utterly incomparable because they are answers to two different questions, making any all-things-considered decision, when these answers conflict, unintelligible. If we decide to act on an agent-neutral reason to violate a conflicting agent-relative requirement, then despite our not having clearly made a wrong decision, our chosen action may be unthinkable. It is with the dirt of acting as an administrator of what ought to happen—as a kind of non-agent, when agency is urgently called for—that we do the unthinkable. The paper concludes that this combination of unintelligibility and unthinkability, when brought on by being an agent who steps outside of our own agency, is the mark of a dirty hands dilemma.

Keywords Dirty hands · Moral dilemma · Moral failure · The unthinkable · Tragedy · Agent relativity

✉ Lisa Tessman
Ltessman@binghamton.edu

[1] Philosophy Department, Binghamton University, Binghamton, NY 13902-6000, USA

1 Introduction

Michael Walzer's "Political Action: The Problem of Dirty Hands" appeared in 1973 against the backdrop of a discussion that had recently begun to emerge about moral conflicts. In 1962 E.J. Lemmon described distinct sources of moral "oughts" (Lemmon 1962: 140), noting that conflicts arise amongst "oughts" with different sources, and concluding that "it is a nasty fact about human life that we sometimes both ought and ought not to do things" (Lemmon 1962: 150). Soon after, in 1965, Bernard Williams published a piece called "Ethical Consistency" that largely shaped the debates about moral dilemmas that have taken place over the decades since. Williams argued that moral conflicts are more like conflicts of desires than like conflicts of beliefs, in that "moral conflicts are neither systematically avoidable, nor all soluble without remainder" (Williams 1965: 117). Add to this Thomas Nagel's (1972) article, "War and Massacre," which Walzer explicitly draws on and which situates the question of unavoidable moral wrongdoing in the context of war and politics. The war in Vietnam, and specifically the My Lai massacre, may have prompted Nagel's thoughts, though he is clear that the massacre was a straightforward wrongdoing that did not present a dilemma because the wrongdoing was avoidable and thus in no sense justified, and he suggests that the entire war in Vietnam was likely unjustifiable as well. However, Nagel notes that decisions that politicians make regarding war more generally are often dilemmatic and he makes the crucial point that they arise because of a conflict between different kinds of moral reasons. When Walzer picks up on the discussion where Nagel left off, he assigns the label of *dirty hands* to a species of dilemma that is endemic to politics, both in the context of war—about which Walzer theorizes extensively in his later work (Walzer 1977, 2004)—and in other political contexts. I agree that dirty hands dilemmas are frequently faced by politicians and that war regularly generates such dilemmas, but, as a moral theorist rather than a political theorist, I would like to follow the clues contained in Nagel's work to understand not only why politicians must often dirty their hands—as identified by Walzer—but also why the problem of dirty hands plagues many people's lives outside of politics.[1] I believe that moral dilemmas are a regular feature of moral life and that the impossible to fulfill demands of morality make us vulnerable to tragic failures and losses. Here I want to consider the features of the subset of moral dilemmas that it is useful to classify as dirty hands dilemmas. By

[1] Walzer does acknowledge that the phenomenon of dirty hands occurs outside of politics: "No doubt we can get our hands dirty in private life also, and sometimes, no doubt, we should" (Walzer 1973: 174); however, his focus–as the title of his article indicates–is on politics. Many authors writing about dirty hands assume that the problem belongs specifically to politics, or that it occurs because of a clash between the different moralities applicable in the public and private realms (see, for instance, early discussions related to the dirty hands literature in Hampshire ed., 1978). A previous collection of responses to Walzer's essay, *Cruelty and Deception: The Controversy Over Dirty Hands in Politics* (Rynard and Shugarman eds., 2000), also situates dirty hands within politics (as the subtitle suggests), though not all the authors in the collection believe that the phenomenon is limited to politics.

borrowing Nagel's ideas about the availability of both agent-relative and agent-neutral perspectives, and about the possibility of conflicts between reasons accessible from these perspectives, and connecting them with my own (Tessman 2015, 2017) account of moral dilemmas and George Harris' (2006) characterization of tragedy, I will propose what I think is one useful way of identifying the distinctive features of dirty hands dilemmas and of the tragic failures and losses that these dilemmas yield.

2 Two Perspectives, Two Questions

Nagel presents the moral dilemmas of warfare as occurring because of the viability of two kinds of moral theories, which tend to produce conflicting directives; using the terms *utilitarianism* and *absolutism* to refer broadly to these two kinds of theories, he characterizes utilitarianism as permitting any means to a good outcome, and absolutism as prohibiting some actions and thus some means. What is most revealing is how he conceives of these two types of theories as having different starting points:

> Utilitarianism gives primacy to a concern with what will *happen.* Absolutism gives primacy to a concern with what one is *doing*. (Nagel 1972: 124)

There are reasons why one thing should happen rather than another, and there are reasons why one should not do certain things. A utilitarian approach recognizes only reasons for what should happen and has nothing to say about what one must do or not do independently of what should happen; for an absolutist, when there exist reasons not to do certain things, these reasons are overriding and so they prohibit one from doing these things even if one's doing these things is part of what should happen. Nagel both endorses absolutist constraints—which allows him to consider how these constraints apply in warfare—and recognizes that even the constraints that he endorses must sometimes be abandoned: "there may be circumstances so extreme that they render an absolutist position untenable" (Nagel 1972: 136). This means that one might have to do something that is terrible and unjustified as a decision about *what one should do*, despite being in some sense the right decision—namely the right decision about *what should happen* (Nagel 1972: 136–137). These two decisions also call for different modes of justifications: the justifications that absolutism requires "are primarily interpersonal" while the justifications that utilitarianism accepts "are primarily administrative" (Nagel 1972: 137–138).

In later work Nagel further develops the idea of there being reasons that are accessible from two different perspectives, eventually (1986) labeling these reasons (borrowing the terminology from Parfit 1984: 143) *agent-relative* and *agent-neutral*

reasons.[2] Nagel emphasizes that these reasons differ in kind precisely because they are accessed from two different *perspectives* or "two ways of looking at the world":

> On the one hand there is the position that one's decisions should be tested ultimately from an external point of view, to which one appears as just one person among others. The question then becomes, 'What would be best? Which of the acts within my power would do the most good, considering matters from out here, impersonally?'…On the other hand there is the position that since an agent lives his [sic.] life from where he is, even if he manages to achieve an impersonal view of his situation, whatever insights result from this detachment need to be made part of a personal view before they can influence decision and action. The pursuit of what seems impersonally best may be an important aspect of individual life, but its place in that life must be determined from a personal standpoint, because life is always the life of a particular person…. (Nagel 1979: 205)

It is when we are in a position to take both of these two different perspectives, asking both "what ought to happen?" and "what ought I do?", that, I believe, there is the risk of facing dirty hands dilemmas. I will come back to this, after some discussion of moral dilemmas more generally.

3 Moral Dilemmas: Tragic Losses and Failures

There are some theorists who have contributed to the discussion of the problem of dirty hands since 1973 who do not consider dirty hands cases to be moral dilemmas, but this is primarily because of the way in which they define moral dilemmas—for instance, they may define them as *irresolvable* moral conflicts (see, e.g., Coady 2018). In contrast, I do not take the issue of resolvability to determine whether a conflict qualifies as a moral dilemma. Instead, I take a moral conflict to be a moral dilemma if, even when resolved, it leaves—to utilize Williams' concept—a

[2] Nagel's terminology evolves. For instance, in "The Fragmentation of Value" he speaks of reasons as having their sources in different kinds of value, calling those that have to do with what happens "impersonal" and "outcome-centered" and "objective" and those that have to do with what one does "personal" and "agent-centered" and "subjective" (Nagel 1979: 133), though he later (1986) uses the terms "objective" and "subjective" differently (and argues that both kinds of reasons can be objective). In "Subjective and Objective" Nagel makes it clear that the reasons are accessible from different points of view (Nagel 1979: 205). In *The View from Nowhere* Nagel defines the terms "agent-relative reason" and "agent-neutral reason": "If a reason can be given a general form which does not include an essential reference to the person who has it, it is an *agent-neutral* reason. For example, if it is a reason for anyone to do or want something, that it would reduce the amount of wretchedness in the world, then that is a neutral reason. If on the other hand the general form of a reason does include an essential reference to the person who has it, it is an *agent-relative* reason. For example if it is a reason for anyone to do or want something that it would be in *his* [sic.] interest, then that is a relative reason" (Nagel 1986: 152–153). He argues that ethics must recognize the legitimacy of both kinds of reasons (and further breaks down agent-relative reasons into three types: reasons of autonomy, deontological reasons, and reasons of obligation), and reiterates his original way of putting it: "Ethics is concerned not only with what should happen, but also independently with what people should or may *do*" (Nagel 1986: 165).

remainder. To be more precise: Moral dilemmas are a subset of moral conflicts. A moral conflict is a situation in which there is a moral requirement to do (or refrain from) one thing and a moral requirement to do (or refrain from) another thing, and in which one cannot do (or refrain from) both. What is distinctive about the subset of moral conflicts that are moral dilemmas is that, in a moral dilemma, neither of the conflicting moral requirements ceases to be a moral requirement as a result of the conflict. One moral requirement may override the other for the purpose of deciding what to do, and when this happens we can say that the conflict has been resolved because one has determined which action is the "right action" in the sense that one knows which action to perform. But in another sense of "right action," if the conflict is a dilemma, then there is no right action, because both the possible actions are wrongdoings. In a moral dilemma, even if one of the conflicting requirements overrides the other, the overridden requirement is not thereby canceled; it remains binding, despite now being impossible to fulfill, and one inevitably fails to fulfill it. The action is a wrongdoing in the sense of being an (unavoidable) violation of a moral requirement. Moral dilemmas thus leave a moral remainder, which may manifest in a number of different emotions, such as guilt, shame, or more generally an anguished sense of responsibility for the failure.[3] What needs some explanation is how there can be moral requirements that are impossible to fulfill—a claim that suggests that "ought" does not always imply "can."

Not *all* overridden moral requirements remain binding despite becoming impossible to fulfill. I distinguish between "negotiable" and "non-negotiable" moral requirements; only non-negotiable moral requirements remain binding no matter what—even if they become impossible to fulfill. A moral requirement is negotiable—and can be canceled—if its associated value can be substituted for, without unique loss, by the value that is associated with the overriding moral requirement; then one might negotiate it away by accepting in its place the fulfillment of the requirement that enables this substitution. No substitution is possible when values are incomparable. A moral requirement can also be canceled if, despite the fact that not fulfilling it will cause unique loss, this loss is the kind of loss that people ought to expect to bear. If a loss is of a sort that we should expect to bear in an ordinary, non-tragic life, then even if, because of incomparability, nothing can exactly substitute for the lost value, the cost can be compensated for by the benefits that depended on that cost. Thus, a moral requirement is negotiable if either substitution or compensation is possible; if neither is possible, then the moral requirement is non-negotiable.

For instance, a moral requirement to protect the life of someone who is in one's care may be a non-negotiable moral requirement, because the unique value of the human life that is lost if one fails to fulfill this requirement cannot be substituted for by any other value, nor, depending on the circumstances of the death, is this the

[3] Stephen De Wijze proposes the term "tragic-remorse" to characterize the emotion indicative of a moral remainder in dirty hands cases (De Wijze 2004). I believe that there is a cluster of emotions that can indicate moral remainders and that which emotion is fitting depends on the details of the situation. I use "anguished sense of responsibility" as a broad category that can be specified in different ways depending on the features of particular cases.

sort of loss that one can necessarily be expected to take in stride, so nothing can compensate for it. In contrast, if someone breaks a promise in order to attend to a much more urgent and important demand—to offer greatly needed help to someone else, for instance—then, if it was a promise regarding something relatively trivial, one should take the broken promise in stride, and will likely feel that the cost to oneself is more than compensated for by the good that it enabled. The line between negotiable and non-negotiable moral requirements depends in part on there being a distinction such as the one that Martha Nussbaum invokes when she describes non-tragic costs as "costs that are to be borne" in contrast to tragic costs that "consist in being made to bear a burden that no [one] should have to bear" (Nussbaum 2000: 1005), because the notion of compensation—which is possible for trivial losses but not for losses of that which makes life worth living—depends on there being some such distinction.

Generally speaking, Kantian or other absolutist approaches take all moral requirements to be non-negotiable, and most consequentialist approaches take all moral requirements to be negotiable. I argue that both these approaches are partly mistaken (and support a pluralist position instead), because neither acknowledge that moral requirements vary in whether or not they are negotiable. Both of these approaches deny the existence of moral dilemmas: Kant denies that duties can conflict, so all moral requirements (all of which he takes to be non-negotiable) can be fulfilled; consequentialism can accommodate conflicts between values, but by taking different values to all be reducible to some common value such as well-being, and by defining right action as the action that maximizes this common value, standard versions of consequentialism always offer a way to resolve conflicts without unique loss: the value associated with the overriding moral requirement is always assumed to substitute for the value that was associated with the overridden (and negotiable, since substitution is possible) moral requirement. But if, contrary to either of these approaches, there are a plurality of kinds of moral requirements—some non-negotiable and some negotiable—and if non-negotiable moral requirements are sometimes overridden in conflicts, then because non-negotiable moral requirements remain binding even if overridden, there can be moral dilemmas.[4]

This line that I am drawing between negotiable and non-negotiable moral requirements is not a clear, bright line, but neither is it meaningless. It is a line that can only be drawn at all because of the human capacity to experience tragedy, that is, because there are tragic losses associated with the violation of non-negotiable moral requirements. In identifying some failures to fulfill (impossible) moral requirements, and the losses associated with these failures, as tragic, I am suggesting something that is in the same spirit as George Harris' (2006) view of tragedy. Harris believes that the capacity for tragedy is what is most distinctive about human beings (Harris 2006: 27–28). Harris assumes, as do I, that values are irreducibly different in kind, that they can conflict, and thus that loss is inevitable. He also distinguishes between different kinds of losses: in some conflicts of values, he points out, one can rationally compare the conflicting values

[4] For a fully developed version of the ideas in the previous four paragraphs, see Tessman (2015, 2017).

and choose to preserve the greater value and sacrifice the lesser; in these cases, although we may experience regret over what is lost—and if the loss is significant this in itself may be tragic—at least we have the "consolation of reason": we know we have rationally chosen what is better, and this is what offers us consolation for the loss. He dubs this "reason's regret" (2006: 30). Tragedy may also take place, of course, if we choose wrongly in such situations, whether by intentionally choosing evil instead of good or whether by mistakenly choosing something of lesser value over something of greater value. But the most deeply tragic cases take place when the conflicting values are incomparable. What I have emphasized about such cases is that nothing can substitute or (if it is also a very significant value) compensate for the sacrificed value, and I have suggested that this is what makes the associated moral requirements non-negotiable; what Harris emphasizes is different but related, namely, that because of incomparability, we can never *know* that we have made the better choice and so we do not even have the consolation of knowing this. He focuses not on the sorts of *moral requirements* that depend on the incomparability of conflicting values (and so not on the existence of the category of moral requirements that I call non-negotiable), but rather on the particular experience of loss to which these conflicts lead:

> Where incomparability obtains and significant values are at stake, plurality and conflict bring loss for which no comparative reason can be given. This sense of loss is not that of regret, nor of waste or consolable loss, but of grief. It is loss without the consolation of reason or the aid of philosophical and religious opiates. This is reason's grief. (Harris 2006: 31)

Reason's grief is "the emotional recognition of unintelligible loss" (Harris 2006: 31). Every value is vulnerable to being tragically and unintelligibly lost, Harris argues, because there is no "unqualified value," by which he means both "that there is nothing of such a value that it cannot in some context be traded off for the value of something else" and "that what is good and what is bad are in an important sense inseparable" and so "there is absolutely no value that is not subject to loss, even given rational choice" (Harris 2006: 29). In some circumstance one might even decide to sacrifice the person one most loves. Harris intends to "haunt" his readers by posing a question that cannot be intelligibly answered: "just how many decent people is your most cherished loved one worth to you?" (Harris 2006: 54). One could never know that any answer to this question is right.

Thus while I might say that tragic loss is the loss of values for which there can be neither substitution nor compensation, and thus that the violation of non-negotiable moral requirements is among the causes of tragic loss, Harris would say (or would add) that in the worst cases—those in which inconsolable grief is warranted—the tragic losses are those in which no consolation is possible because incomparability has made a significant loss unintelligible. Harris' account sheds light on the specific sort of tragedy brought on by dirty hands dilemmas, because, as I will argue, unintelligibility turns out to be an element of dirty hands cases. The account that I have (in previous work) offered of the experience of violating a non-negotiable moral requirement focused not on the experience of unintelligibility, but rather on the experience of unthinkability, which, I will suggest, is another element of dirty

hands dilemmas. That is, I believe that my account and Harris' are compatible, and that both unintelligibility and unthinkability play a role in dirty hands dilemmas.

People have the experience of unthinkability when they judge themselves to be absolutely required to do something, even if doing it is impossible: it is simply unthinkable not to do it. Thus it is important to understand how people make this sort of judgment. A judgment that one is absolutely, though impossibly, required to do something is an intuitive moral judgment, made through a cognitive process that does not involve conscious reasoning. While we may utilize the principle that "ought implies can" when we consciously reason about what we are morally required to do, this principle may not affect our affect-laden, intuitive moral judgments. Moral reasoning in accordance with deontic logic explicitly includes the principle that "ought implies can" as a premise, thus ruling out the conclusion that one is impossibly required to do something. Consequentialist reasoning is less explicit about its use of "ought implies can," but because consequentialist reasoning involves choosing the best *possible* action—defined, for instance, as the action that will result in the best possible state of affairs—in any situation, actions that are impossible to perform are not even considered. It is when neither form of reasoning is employed that we can reach the verdict—through an automatic, unconscious process—that what we must do is something that we cannot do. We might simply find ourselves with the verdict: "I must do this" or "I must not do this," or even "doing *that* would be unthinkable." When we intuitively take some action to be unthinkable, and in that sense we experience it as impossible to perform, then even if its alternative is also impossible (in a more literal sense), we experience the impossible alternative as absolutely required. For instance, consider a situation in which one must betray—fail to protect—one of one's two children, as the fictional character Sophie was forced to in *Sophie's Choice* (Styron 1976). Because the only way to avoid betraying and sacrificing one child is by betraying and sacrificing the other child—which is unthinkable—each of the two available actions is both unthinkable and, because it is the only route to avoiding the equally unthinkable alternative, absolutely required. One cannot perform both of the two absolutely required actions, so the result is inevitable failure to fulfill one of the requirements. The unthinkability of each action automatically triggers the judgment that one is impossibly required to avoid it.[5]

Nagel's insight that there are "two ways of looking at the world" (Nagel 1979: 205) can help us understand how we find ourselves facing impossible moral requirements. We only see moral requirements as non-negotiable when we take an agent-relative perspective. From an agent-neutral perspective, values appear to be comparable and substitution appears to be possible, and thus no moral requirements seem to be non-negotiable. Having the experience of unthinkability thus depends on whether we take an agent-relative perspective or an agent-neutral perspective: we only experience unthinkability when we ask the question of what we should do. The experience of unthinkability is a subjective experience, directly accessible only to the person whose experience it is. What is unthinkable about committing an act that is absolutely prohibited—an act of murder, or torture, for instance—is *being the*

[5] Full discussion of the ideas in this paragraph can be found in Tessman (2015, 2017).

person who commits that act. Thus, the unthinkability of an act cannot be grasped from an agent-neutral position. A moral requirement may be non-negotiable when viewed from an agent-relative perspective, but we will not *experience* violating it as unthinkable if we do not judge it from an agent-relative perspective.[6]

In a case in which the *only* perspective that one takes is agent-relative, it is possible for an agent to experience both of two conflicting alternatives as unthinkable. If one is forced to choose which of two children to betray and sacrifice, and one remains within an agent-relative perspective, asking "what should I do?" while making judgments about each option, then one will intuitively judge each possible action to be unthinkable (non-negotiably prohibited) and non-negotiably required (because if one does not perform that action, one must carry out the only alternative, which is also non-negotiably prohibited). Neither the agent-relative reason to protect one child nor the agent-relative reason to protect the other child can be overriding, and one must simply pick one in the absence of an overriding reason. If, instead, the situation were to prompt one to take *both* an agent-relative and an agent-neutral perspective, one would additionally ask "what would be best?"; one could from this additional agent-neutral perspective consider, for instance, whether one child or the other would have a greater chance of survival, and then act on this agent-neutral reason (and as I will argue, it is this that would entail dirtying one's hands).

In other kinds of cases, one action seems to be morally required when one asks, from an agent-relative perspective, what one should do, and the conflicting alternative seems to be morally required when one asks, from an agent-neutral perspective, what should happen. We are not psychologically prone to making the same sort of affect-laden, intuitive judgments when we occupy an agent-neutral perspective, so we would in this kind of case experience only one of the two alternatives as unthinkable. Nevertheless, both options could still be considered to be non-negotiably required (or prohibited), in the sense that *if* we were to view each option from an agent-relative position, we would see each as involving the sacrifice of values for which there can be neither substitutions nor compensation.

Consider, for instance, a version of trolley dilemma (the "push" or "footbridge" version) in which one must either push someone to their death so that their body will serve as a trolley-stopper or else allow the trolley to kill five other people trapped on the tracks. Most people take only an agent-relative perspective when presented with this emotionally charged version of the dilemma. They then ask what they should do in the situation, and intuitively judge it to be unthinkable to push someone to

[6] There may also be other causes of our not finding it to be unthinkable to violate a non-negotiable moral requirement. For instance, even if I take an agent-relative perspective and ask how I ought to treat a person who is right in front of me, if racism leads me to devalue that person, perhaps thinking of them as sub-human, I may find it acceptable to do terrible things to them. However, because the construction of moral requirements is a social process, whether an act is required, permitted, or prohibited–and whether this is negotiable or non-negotiable–is not determined by any one person's subjective experience. Thus we might say that an act is (i.e. has been socially constructed as) non-negotiably prohibited, even if only some people have the experience of its unthinkability, and even if having this experience varies depending on whether someone adopts an agent-neutral or agent-relative perspective. For more on the construction of moral requirements, see Tessman (2015, 2017).

their death. A very small minority of respondents take an agent-neutral perspective instead and ask only what should happen, concluding that the one person should be pushed in order for the five to be saved; however, this small minority of respondents do not make their decision intuitively and thus they avoid experiencing unthinkability as part of the process of arriving at this decision. Nevertheless, each alternative does involve the loss of one or more human lives, which is surely the sort of value for which (as becomes apparent when one views the situation from an agent-relative perspective) there can be neither substitutions nor compensation. So there is a sense in which we ought to find each choice to be both non-negotiably prohibited and non-negotiably required (just as someone does when trying to choose, from a solely agent-relative perspective, which of their two children to sacrifice). The situation is actually a dilemma, a case of unavoidable moral failure. Taking refuge in an agent-neutral perspective allows one to avoid experiencing it as such, for when we take *only* an agent-neutral perspective, we do not experience any choice as dilemmatic or as involving tragic loss: the answer to the question, “what should happen?” is always “whatever is best,” and arriving at this answer requires a presumption of comparability.

We must conceive of the occupant of an agent-neutral perspective as a sort of non-agent, removed from all particular, interpersonal relationships and thus from interpersonal practices of accountability. To a non-agent, nothing is valued in its particularity because only a valuing creature can actually *value*. The occupant of an agent-neutral perspective is not, *qua* this sort of non-agent, a valuing creature.[7] At best, the occupant of an agent-neutral perspective surveys the values of all the relevant valuing creatures and abstracts from these to maximize overall value, and the abstracted values appear comparable in a way that actual values are not. Harris might make the point here that the decision will then seem intelligible. I would add a point about unthinkability: if one occupies *only* an agent-neutral perspective when thinking about the conflict, whatever action resolves the conflict for the best will simply seem to be right rather than wrong, let alone unthinkable.

Dirty hands dilemmas occur only when we are thrust into both of the two perspectives, agent-relative and agent-neutral, and face two questions at once: the question of what we should do and the question of what should happen. If we ask *only* the agent-relative question of what we should do, we might still find ourselves with a dilemmatic choice, and though this can be tragic and warrant “reason’s grief,” we do not dirty our hands in making or acting on our choice. If we ask *only* the agent-neutral question of what should happen, we face no dilemma at all. But answering the two questions at once can lead us to dirty our hands. Answering both questions involves comparing what is especially incomparable, and this gives rise to an especially acute kind of unintelligibility; and, because what we unintelligibly decide to do in dirty hands cases is to violate the non-negotiable moral requirement that is supported by reasons accessible from an agent-relative perspective, we also

[7] Technically, I want to allow that a valuing creature might not be an agent. I don’t want to take up here the question of what is necessary for agency, but I will assume that there may be valuing creatures, human or not, who would not count as agents.

experience this as an unthinkable moral failure. It is with all this in mind that I will propose what I take to be the conditions of dirty hands dilemmas.

4 Dirty Hands: A Collision of Two Questions, an Unintelligible and Unthinkable Resolution

Walzer's groundbreaking article on dirty hands suggests that political leaders are in a special position that is characterized by a high risk of having to dirty their hands, and he aims to pinpoint what exactly it is about politics that makes it so likely a realm for dirty hands. He adopts from Nagel the framework of the opposition between utilitarianism and absolutism, and he takes the implication of the opposition to be that, because the *role* of a political leader "imposes on him [sic.] a considerable responsibility for consequences and outcomes" (Walzer 1973: 161), political leaders can expect that they may have to navigate the opposition by violating absolutist constraints for the sake of doing what is right according to the utilitarian calculation. Beyond this, Walzer argues that the problem of dirty hands "is posed most dramatically in politics for the three reasons that make political life the kind of life it is, because we claim to act for others but also serve ourselves, rule over others, and use violence against them" (Walzer 1973: 174). While Walzer's portrayal of the position that political leaders are in and of the demands of their office depends on the assumption that utilitarian and absolutist requirements can conflict in ways that make it impossible to "govern innocently" (to quote Hoerderer's phrase from Sartre's 1948 *Les Mains Sales* [Sartre 1989]), he leaves in the background Nagel's focus on the *root* of the differences between these two moral theories: the "distinction between what one does to people and what merely happens to them as a result of what one does" (Nagel 1972: 131), or, put differently, the distinction between taking an agent-relative perspective and taking an agent-neutral perspective. In my view, this distinction is key to understanding dirty hands dilemmas, and the reason that it is so crucial is that it is as a result of having to take *both* perspectives—thus being an agent who meanwhile renounces the interpersonal responsibilities of agency—that we risk dirtying our hands. I want to bring this distinction to the forefront in order to see how being both an agent and a sort of non-agent at the same time invites a particularly insuperable kind of unintelligibility and does so in a situation in which we cannot avoid the subjective experience of unthinkability. It is this combination of unintelligibility and unthinkability, brought on by stepping outside of an agent-relative perspective, that I take to be the mark of dirty hands dilemmas.

Here, then, are my proposed criteria for dirty hands dilemmas. One is in a dirty hands dilemma when one faces a dilemma that one resolves by dirtying one's hands. This happens when the following conditions are met:

1. One takes both an agent-relative perspective, asking "what ought I do?" and an agent-neutral perspective, asking "what ought to happen?"
2. The answer to "what ought to happen?" can only be actualized if one does A, but the answer to "what ought I do?" is that one ought not do A.

3. One decides that the action-guidance offered by the answer to “what ought to happen?” overrides the action-guidance offered by the answer to “what ought I do?”, and so one does A; and, this decision is not clearly a wrong decision.
4. The two answers—because they are answers to two different questions—are utterly incomparable, and thus the decision about which action-guidance is overriding is unintelligible.
5. The action that one takes (namely A) is unthinkable (and non-negotiably prohibited) from an agent-relative perspective.

A dirty hands dilemma is a dilemma that one navigates by doing something that is both unintelligible and unthinkable, but that one nevertheless gets through without making what is clearly a wrong decision; the experience of dirtying one’s hands is thus an experience of both tragic loss that warrants inconsolable grief, and of tragic, unavoidable moral failure.

Each of the five conditions call for elaboration.

1. One takes both an agent-relative perspective, asking “what ought I do?” and an agent-neutral perspective, asking “what ought to happen?”

The first condition—suggested by Nagel’s distinction—is what makes politics an obvious location for dirty hands dilemmas, as long as one conceives of political leaders in their role as administrators of “what ought to happen,” for this would make them likely candidates for taking both an agent-relative and an agent-neutral perspective. Other people may, in their ordinary, personal lives, not have many occasions to take both perspectives; much of the time they may be able to occupy a solely agent-relative perspective, making decisions about what to do and being interpersonally accountable to others for their decisions and actions. This does not guarantee that they will not encounter dilemmas, for it is possible for responsibilities to others to conflict and for these conflicts to be resolvable only with remainder.[8] However, I would not count such dilemmas as dirty hands dilemmas if the person facing the dilemma never adopts an agent-neutral perspective. For instance, when I discussed the case, modeled on *Sophie’s Choice*, of choosing which of one’s two children to betray and sacrifice, I imagined someone like Sophie in two different ways. One possibility is that she remains solely within an agent-relative perspective and can find no agent-relative reason to support one action over the other; she nevertheless must act, so she picks one (sacrificing her daughter) and does the unthinkable. This is an

[8] In fact, Christopher Gowans’ entire argument for the conclusion that there are situations of inescapable moral wrongdoing is built on the claim that it is when moral values of a particular kind, which he calls “responsibilities to persons,” come into conflict with each other that moral wrongdoing may become inescapable (Gowans 1994: 121). He defines “responsibilities to persons” in such a way that it is clear that they are agent-relative moral requirements (Gowans 1994: 121–128). While I agree with Gowans that conflicts between responsibilities to persons may leave moral remainders, I am arguing that other kinds of moral conflicts may also leave moral remainders; in dirty hands cases, the conflict must be between a moral requirement that is supported by an agent-relative reason and a moral requirement that is supported by an agent-neutral reason.

example of unavoidable moral failure, but it is not a dirty hands case. The other way to imagine Sophie—and it is in this scenario only that I would say that she dirties her hands—is to picture her as taking both an agent-relative and an agent-neutral perspective when she is forced to choose between her children. Then she might find an agent-neutral reason that would support one choice over the other, such as "the boy is older and has a greater chance of survival, so choosing him would increase the chance of at least one child surviving, which would be best." My claim is that Sophie, *only* when imagined in this way, dirties her hands precisely by shifting into an agent-neutral perspective in the very moment in which her agency is crucial for the possibility of her remaining bound by her agent-relative responsibilities toward each of her children (though of course the tragedy is that she is unable to carry out these responsibilities toward one of them even if she maintains an agent-relative perspective). The moral failure associated with a dirty hands dilemma is not worse than the moral failure associated with a dilemma in which one does not dirty one's hands—my claim is merely that it is distinctive: we dirty our hands only by acting on agent-neutral reasons for violating agent-relative requirements. In contrast, in dilemmas that are not dirty hands cases, we violate one agent-relative requirement by acting on another, conflicting agent-relative requirement.[9]

Here I am agreeing with Michael Stocker's claim that not "every conflict involving wrong acts involve[s] dirty hands" (Stocker 1990: 10), but I disagree with him about why. For Stocker, although dirty hands cases extend beyond politics, they are distinguished from other cases of justified or obligatory wrongdoing by the fact that paradigmatic dirty hands cases take place in a context of immorality, in which the agent is coerced into participating in the immorality (Stocker 1990: 20). He contrasts Sophie's situation with a situation in which a "natural calamity" causes one to have to choose which child to keep alive, claiming that only cases like Sophie's, which involve forced participation in an immoral project, are dirty hands cases (Stocker 1990: 19–20). I believe that my two ways of imagining Sophie get at a feature (namely, stepping into an agent-neutral perspective while also continuing to occupy an agent-relative perspective) that is the best marker of the phenomenon of dirty hands. One implication of my first condition is that some cases of coerced participation in immorality are not dirty hands cases (such as when Sophie never takes up an agent-neutral perspective), and some dilemmas that do not involve immorality in this way (such as when one does take up an agent-neutral perspective to decide what ought to happen in the face of a natural disaster) are dirty hands dilemmas.

[9] Nick (2019) has argued in favor of what she calls a "symmetrical" view rather than an "asymmetrical" view of dirty hands dilemmas, and according to her definitions, she would count my view as an asymmetrical view because it counts acting on agent-neutral reasons when they conflict with agent-relative reasons as dirtying one's hands but does not count acting on agent-relative reasons in moral conflicts as dirtying one's hands. However, her reason for rejecting asymmetrical views is that she takes these views to come with a cost: the cost of not recognizing that moral failure is unavoidable in all moral dilemmas, regardless of what kind of values conflict and regardless of which "horn" one chooses. My view does not come with this cost, because I count all moral dilemmas as situations of unavoidable moral failure. Keeping one's hands clean does not imply avoiding moral failure. In my view, dirtying one's hands is just one distinctive way of failing.

I believe it is simply being moved out of one's own agency—into an agent-neutral position—and making a judgment from the perspective of a sort of non-agent, that is a condition of dirtying one's hands. It is, one might say, a context of administration, rather than necessarily a context of immorality, that is required for dirty hands. The dirt is not that of immorality, but rather of a sort of non-agency when interpersonal agency is still called for, and when interpersonal agency is what *should* (from an agent-relative perspective) still stop us from doing the terrible thing that we do.[10] So I will say that when one retains an agent-relative perspective but also adopts an agent-neutral perspective, the first condition for dirtying one's hands is met.

Is politics a realm in which this condition is especially likely to be met? As Nagel puts it, the agent-neutral perspective of utilitarianism "is associated with a view of oneself as a benevolent bureaucrat distributing such benefits as one can control to countless other beings, with whom one may have various relations or none" (Nagel 1972: 137–138). It is true that some of what political leaders do is to make decisions with a view to achieving some overall good. They are in a context of administration. But I have two caveats. The first, which I will just mention but not develop, is that this is an impoverished view of politics: if one conceives of politics (perhaps along Arendtian lines) as contestation rather than bureaucratic administration, one would not see politics as such an obvious site for dirty hands dilemmas. The second is that political leaders are far from being the only people whose formal role involves this kind of administration, and still others without any formal role may find themselves positioned to administer costs and benefits. For instance, people who serve as university administrators or administrators in health care organizations regularly have the responsibility of determining what ought to happen, or what would be best, within their institutions, and sometimes the answers to these questions do involve sacrificing some people's rights or lives (or violating other absolutist requirements) for the sake of some other good. In health care there are clearly life-and-death decisions to be made, such as decisions about how to ration scarce resources.[11] Even

[10] My account of dirty hands not only departs from Stocker's, it also departs from those who adopt and further develop Stocker's account, such as De Wijze (1994, 2002). De Wijze is more explicit than Stocker is about what the dirty feature is: "in all cases of dirty hands what is common is that actions involve the *justified* betrayal of persons, values or principles due to the immoral circumstances created by other persons (or organization of persons) within which an agent finds herself… How is this situation different from ordinary cases of horrendous moral conflict or dilemma? The difference lies in who/what created the evil circumstances. Dirty hands occur when the evil circumstances are created by other human beings…Evil human agency is crucial to the dirty hand scenario" (De Wijze 1994: 30). Elsewhere De Wijze defines evil by drawing on intuitions about what the "dirt" is in dirty hands cases, noting that "the 'dirt' that adheres to (and is felt by) persons in such situations provides insights into the uniqueness of evil and the nature of its moral residue" (De Wijze 2002: 212). I have different intuitions than both Stocker and De Wijze do about which sample cases count as examples of dirty hands; for Stocker's sample cases, see (Stocker 1990: 24), and for De Wijze's sample cases, see (De Wijze 1994: 30–31).

[11] One of Phillip Tetlock and colleagues' experiments involves having subjects read a story about a hospital administrator who faces what they call a "taboo trade-off"; he is in a position to decide–from an agent-neutral position–"what will be best," while also being confronted with an agent-relative reason to save a particular child's life. The story about the hospital administrator is as follows:

Robert can save the life of Johnny, a five year old who needs a liver transplant, but the transplant procedure will cost the hospital $1,000,000 that could be spent in other ways, such as purchasing better equipment and enhancing salaries to recruit talented doctors to the hospital. Johnny is very ill and has been on

people who do not determine the administrative or institutional policies that affect them at least must interpret these policies or decide whether to actually abide by them. The soldier occupies both an agent-relative and agent-neutral perspective in deciding whether to shoot or not shoot at an "enemy" child who is about to throw an explosive at their unit's encampment; the soldier asks both "is it morally permissible for me to kill a child?" (no) and "what would be the best possible outcome of this situation?" (that the child be shot and the unit saved). The kindergarten teacher occupies both an agent-relative and agent-neutral perspective in deciding whether to risk the safety of the whole class of children by opening the classroom door to the one child who had been in the bathroom when an active shooter entered the building; the teacher asks both "ought I protect this one child who is in my care?" (yes) and "what will minimize the number of children who are killed?" (that the door be barred, thus saving the 24 children hiding in the room). Even the ordinary resident of a neighborhood that is flooded by a storm can find themselves suddenly taking both an agent-relative and agent-neutral perspective if they spot a rowboat that they can commandeer and must decide exactly what to use it for; they ask both "is it permissible for me to fight off people who are trying to stay alive by clambering onto this boat?" (no) and "what is the optimal number of people to have on this boat, given the risks of it sinking if overloaded and the benefits of getting more people to safety?" (five, and not the added sixth person trying to get on).

Walzer does offer reasons for thinking that it is not simply being in an administrative role that makes political leaders particularly prone to dirtying their hands. Having the power of political office and the ability to enforce one's will, including through violence, facilitate the performance of the kind of acts that violate what, from an agent-relative position, can be seen to be non-negotiable moral requirements: political leaders may have the power to carry out unthinkable acts that other

Footnote 11 (continued)

the waiting list for a transplant but because of the shortage of local organ donors, obtaining a liver will be expensive. Robert could save Johnny's life, or he could use the $1,000,000 for other hospital needs (Tetlock et al. 2000: 858).

If subjects are told that Robert uses the money for "other hospital needs," then they express intense moral outrage at him. In other words, the outrage is a response to the hospital administrator taking the moral requirement that is supported by an agent-neutral reason (considering what best satisfies the hospital's needs, overall) to override the moral requirement that is supported by an agent-relative reason (saving the life of the particular child, Johnny). Additionally, subjects who are told that Robert decides "after much time, thought, and contemplation" to save Johnny's life also express moral outrage about him, but subjects who are told that Robert is very quick to make the decision to save Johnny's life do not express outrage. The outrage at the length of time that Robert takes to act on the agent-relative reason is a response to Robert even *considering* doing the unthinkable by violating a "sacred value"–namely, the value of the particular human life as seen from an agent-relative perspective (Tetlock et al. 2000: 858–859). The outrage at Robert acting–or even considering acting–on the agent-neutral reason was found to be accompanied by subjects' need for "moral cleansing," suggesting that "merely contemplating taboo trade-offs may be sufficient to create a sense of moral contamination (feeling dirty, befouled) that people try to eliminate by strenuously reaffirming their commitment to defending the moral order…" (Tetlock et al. 2000: 860). I take Tetlock et al.'s analysis–and their finding that people feel dirtied by taboo trade-offs–to support my claim that dirty hands cases are those in which someone in a context of administration acts on what they take to be an overriding agent-neutral reason, violating a moral requirement that is supported by an agent-relative reason.

people would not be in a position to carry out. Thus political leaders *are* especially vulnerable to dirtying their hands, but are far from being the only people who are in this position.

Having hands is, metaphorically, what it is to be an agent, to be someone who is interpersonally accountable to others for actions.[12] We dirty our hands by transgressing our responsibilities (*qua* agent) through the very act of relinquishing interpersonal agency, because it is in shedding this agency and occupying an agent-neutral perspective that we find ourselves with a reason to violate agent-relative moral requirements. It is our capacity to enter the space of non-agency that makes us vulnerable to failures in a specific way that is best marked off by saying that we dirty our hands. We are an agent who acts as a non-agent when our interpersonal agency is still absolutely required, because this agency is what we depend on to give us a non-negotiable reason not to violate absolutist constraints.

2. The answer to "what ought to happen?" can only be actualized if one does A, but the answer to "what ought I do?" is that one ought not do A.

The second condition is simply the fact of moral conflict, and the specification that in a dirty hands case, the conflict is between two moral requirements, one of which is supported by an agent-neutral reason and the other of which is supported by an agent-relative reason. This condition does not entail that the conflict is a dilemma, because it does not specify that the resolution leaves a remainder (for that we will need to add, as condition five does add, that the overridden requirement is non-negotiable and thus is not canceled by being overridden). There are, of course, plenty of situations in which agent-relative and agent-neutral reasons point in the *same* direction, permitting one to avoid committing any wrongdoing at all; such situations do not present conflicts. For instance, the doctor who has only one patient in need of care has both an agent-relative and an agent-neutral reason to care for that one patient. In this case, the second condition would not be met.

3. One decides that the action-guidance offered by the answer to "what ought to happen?" overrides the action-guidance offered by the answer to "what ought I do?", and so one does A; and, this decision is not clearly a wrong decision.

The third condition suggests that we only dirty our hands when we act on our answer to the question, "what should happen?" rather than on our answer to the

[12] As Macbeth cries, in his guilty and anguished state, "What hands are here?" Macbeth, though, bloodies rather than dirties his hands. If dirt is not the dirt of immorality, but rather the dirt we step into by entering a space of non-agency, then perhaps blood is what we get on our hands when we straightforwardly violate absolutist constraints as a moral agent. [For a different, and much more developed, distinction between dirt and blood, see Thalos (2018). She argues that at least some cases of dirty hands occur when one person authorizes another agent to commit a morally required wrongdoing; the dirt is the residue from the wrongdoing, and the dirt attaches to the authorized agent; however, in Thalos' view, blood represents the responsibility for the wrongdoing, and this belongs to those who authorized the agent, such as citizens who have authorized a political office holder to act (Thalos 2018: 175–176)].

question, "what ought I do?" when they conflict. This is most commonly characterized as taking place when we reach a certain threshold: all but the very most strict deontologists eventually, as the gloating consequentialists like to point out, cave in and act like consequentialists. Walzer's original example of torturing a terrorist to learn where the bombs are hidden invokes the idea of a threshold being reached—presumably many innocent people will be killed if the terrorist is not tortured (Walzer 1973: 166–167).[13] In later work, Walzer explores where this threshold is by asking what constitutes enough of a "supreme emergency" to warrant violating absolutist constraints by bombing innocent people (Walzer 1977: Chap. 16). Walzer sets the threshold high:

> A morally strong leader is someone who understands why it is wrong to kill the innocent and refuses to do so, refuses again and again, until the heavens are about to fall. (Walzer 2004: 45)

He has something specific in mind for the falling of the heavens:

> Dirty hands aren't permissible (or necessary) when anything less than the ongoingness of the community is at stake, or when the danger that we face is anything less than communal death. (Walzer 2004: 46)

The fact that political leaders serve in a role in which they can expect to have to determine "what ought to happen" means that they face less of a barrier to acting on agent-neutral reasons than do people who never expected to be in a position to decide—especially on a large scale—what ought to happen. People who never intended to have this kind of decision-making power might resist it, and thus resist shifting into an agent-neutral perspective, perhaps with the thought, "I don't want to play God." Those who hold political office may slip more readily into an agent-neutral perspective if they simply assume it comes with their position, and—having crossed that barrier—only need to determine the threshold at which agent-neutral reasons override agent-relative reasons.

If it were possible to know that the decisions about which reasons are overriding, and which action to take, were *wrong* decisions, then the agent who makes a wrong decision would not be said to have dirty hands, but rather to simply have chosen wrongly in a straightforward way. Thus for this third condition to be one of the necessary conditions for dirty hands, it must stipulate that it is not clear that one has decided wrongly (though note that one still commits a wrongdoing even if one has not made a wrong action-guiding decision). However, it stops short of saying that one must make a *right* decision because, as the next condition states, the decision is unintelligible. It is a decision that we cannot know is right and cannot know is wrong. However, in extreme cases we can proceed as if we do know, because

[13] "Threshold deontologists" may say that it is *right* to ignore deontic constraints when the threshold is reached. My position is different: while it may be a correct action-guiding decision to violate deontic constraints in such a case, this violation is not a right act, but rather an unavoidable wrongdoing. Walzer, too, describes the violation of absolutist constraints as wrong, but still to be done when a threshold is reached.

when the heavens are about to fall, as Walzer notes, "no great precision is required" for us to act as if we know that the threshold has been reached: "we can only be overwhelmed by supreme emergency" (Walzer 2004: 40). Thus the decision referred to in condition three is a decision that we may *take* to be right rather than wrong, though it is actually unintelligible.

4. The two answers—because they are answers to two different questions—are utterly incomparable, and thus the decision about which action-guidance is overriding is unintelligible.

The sort of unintelligibility that Harris analyzes occurs because of an incomparability of values. I believe that the unintelligibility of the decision that has to be made in a dirty hands dilemma is even more serious: one is forced to ask two completely different questions and take both answers into account, somehow weighing them against each other if they point toward different actions. That is, in dirty hands cases, one does not just have to compare two different values from the same perspective; one must compare the reasons accessible from one perspective to the reasons accessible from another perspective, and determine which are overriding. Asking oneself, "Would it be best if one terrorist is tortured and the city is saved from the ticking bomb, or am I morally forbidden to torture?" is no more intelligible than asking, "Is chocolate objectively better than vanilla, or do I prefer vanilla?" If chocolate is objectively better than vanilla, *and* I prefer vanilla, I cannot intelligibly know how to answer the disjunctive question. The disjunction in the question must be exclusive if the question is being asked in order to identify which single action (e.g. either ordering chocolate ice cream or vanilla ice cream) is supported by an overriding reason: answering by saying "both" does not yield action-guidance. Similarly, if it is best for one terrorist to be tortured and the city saved, and it is also the case that I am morally forbidden to torture, then there is no intelligible way to respond to the question "Would it be best if one terrorist is tortured and the city is saved from the ticking bomb, or am I morally forbidden to torture?" For the purpose of action guidance, one cannot answer "both," but one also cannot intelligibly arrive at any other answer. If, as Harris suggests, grief is "the emotional recognition of unintelligible loss" (Harris 2006: 31), then the tragic loss associated with the overridden moral requirement in a dirty hands dilemma warrants inconsolable grief. One lacks the "consolation of reason," of knowing that the choice was for the best.

5. The action that one takes (namely A) is unthinkable (and non-negotiably prohibited) from an agent-relative perspective.

Not only does the person who dirties their hands experience an unintelligible, tragic loss for which they are responsible, they also experience the unthinkability of the action that they must take. Again, this follows from taking two perspectives at once: as a non-agent, one accesses the agent-neutral reasons that one takes to be overriding, which entails the decision to violate absolutist constraints; but as an agent who occupies an agent-relative perspective one takes the violation of

these constraints to be unthinkable. As an agent, one is forbidden from even *considering* the violation—that is what it means for it to be unthinkable. The relinquishing of interpersonal agency, then, dirties one's hands because it constitutes an abandonment of the agent's responsibility to never even consider doing the unthinkable.

5 Conclusion

There is much that is unthinkable for us, as human agents who have the capacity to experience tragic loss and grief, and to experience ourselves as being required to avoid the unavoidable moral failures that may bring about such loss. But nothing is unthinkable from an agent-neutral perspective. It is in recognizing that we can occupy both perspectives at once that we can make sense of what it is to think the unthinkable: the non-agent thinks what is unthinkable to the agent. The phenomenon of dirty hands—in which we not only think but also do the unthinkable—appears paradoxical, then, in a way that is similar to other phenomena that depend on taking two perspectives, such as the phenomena of bearing the unbearable, surviving the unsurvivable, and forgiving the unforgivable.

We bear the unbearable when, as an embodied being, we experience severe and prolonged pain as unbearable, and yet we bear it through a kind of transcendence in which we observe our body as if from outside of it. We survive what we experience as unsurvivable through the sheer objectivity of time advancing regardless of our subjective feeling that we cannot go on for one more moment. There is even a way to forgive the unforgivable by shifting perspective: as a person who has been wronged we may find an offense to be unforgivable; but we can take a different perspective, and access, as Cheshire Calhoun has put it, a biographical rather than a moral view of the perpetrator, understanding what they have done as forgivable by choosing "to place respecting another's way of making sense of her life before resentfully enforcing moral standards" (Calhoun 2016: 245).

These are some of the paradoxes that come with the human ability to shift perspective, simultaneously occupying and leaving behind the particular self that we are. We have hands that can be dirtied and they are the hands of a self who can cause and experience unintelligible, tragic loss and who can—sometimes unavoidably—fail to meet interpersonal responsibilities, including in ways that are unthinkable. But nevertheless, from a perspective far outside of this self, that may be exactly what ought to happen.

Acknowledgements Thanks to Stephen De Wijze and Christina Nick for helpful comments on a draft of this paper.

Declarations

Conflict of interest The author does not have any relevant financial or non-financial interests to disclose. There are no relevant competing interests, and there are no conflicts of interest.

References

Calhoun, Cheshire. 2016. Changing one's heart. In *Moral aims: Essays on the importance of getting it right and practicing morality with others*. New York: Oxford University Press.
Coady, C.A.J. 2018. The problem of dirty hands. *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. https://plato.stanford.edu/archives/fall2018/entries/dirty-hands/. Accessed 22 August 2022.
De Wijze, Stephen. 1994. Dirty hands—Doing wrong to do right. *South African Journal of Philosophy* 13(1): 27–33.
De Wijze, Stephen. 2002. Defining evil: Insights from the problem of "dirty hands." *The Monist* 85(2): 210–238.
De Wijze, Stephen. 2004. Tragic-remorse—The anguish of dirty hands. *Ethical Theory and Moral Practice* 7: 453–471.
Gowans, Christopher. 1994. *Innocence lost: An examination of inescapable moral wrongdoing*. New York: Oxford University Press.
Hampshire, Stuart, ed. 1978. *Public and private morality*. Cambridge: Cambridge University Press.
Harris, George. 2006. *Reason's grief: An essay on tragedy and value*. Cambridge: Cambridge University Press.
Lemmon, E. J. 1962. Moral dilemmas. *The Philosophical Review* 71(2): 139–158.
Nagel, Thomas. 1972. War and massacre. *Philosophy & Public Affairs* 1(2): 123–144.
Nagel, Thomas. 1979. *Mortal questions*. Cambridge: Cambridge University Press.
Nagel, Thomas. 1986. *The view from nowhere*. New York: Oxford University Press.
Nick, Christina. 2019. Can our hands stay clean? *Ethical Theory and Moral Practice* 22: 925–940.
Nussbaum, Martha. 2000. The costs of tragedy: Some moral limits of cost-benefit analysis. *Journal of Legal Studies* 29(2): 1005–1036.
Rynard, Paul, and David Shugarman, eds. 2000. *Cruelty and deception: The controversy over dirty hands in politics*. New York: Broadview Press.
Sartre, Jean-Paul. 1989. [1948]. Dirty hands (Les mains sales). In *No exit and three other plays*, New York: Vintage Books.
Stocker, Michael. 1990. *Plural and conflicting values*. New York: Oxford University Press.
Styron, William. 1976. *Sophie's choice*. New York: Random House.
Tessman, Lisa. 2015. *Moral failure: On the impossible demands of morality*. New York: Oxford University Press.
Tessman, Lisa. 2017. *When doing the right thing is impossible*. New York: Oxford University Press.
Tetlock, Philip, Orie Kristel, S. Beth Elson, Melanie Green, and Jennifer Lerner. 2000. The psychology of the unthinkable: Taboo trade-offs, forbidden base rates, and heretical counterfactuals. *Journal of Personality and Social Psychology* 78(5): 853–870.
Thalos, Mariam. 2018. Dirty hands: The phenomenology of acting as an authorized agent. *The Monist* 101: 170–186.
Walzer, Michael. 1973. Political action: The problem of dirty hands. *Philosophy & Public Affairs* 2(2): 160–180.
Walzer, Michael. 1977. *Just and unjust wars: a moral argument with historical illustrations*. New York: Basic Books.
Walzer, Michael. 2004. Emergency ethics. In *Arguing about war*. New Haven: Yale University Press.
Williams, Bernard. 1965. Ethical consistency. *Proceedings of the Aristotelian Society* 39: 103–124.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.