

# **Undercover Policing and ‘Dirty Hands’: The Case of Legal Entrapment**

## **Abstract**

Under a ‘dirty hands’ model of undercover policing, it involves situations where whatever the law-enforcement agent does is morally problematic. Christopher Nathan’s criticism of the model assumes the view that morally wrongful acts are central to undercover policing. We address this criticism, and other aspects of Nathan’s discussion of the ‘dirty hands’ model, specifically in relation to legal entrapment to commit a crime. We explain three dilemmatic versions of the ‘dirty hands’ model. We show that, while two of these are inapplicable to legal entrapment, the third has better prospects. We then argue that, since the third model precludes Nathan’s criticism, a viable ‘dirty hands’ model of legal entrapment remains an open possibility. Finally, we summarize and generalize this result.

**Keywords** Dirty Hands • Legal Entrapment • Moral Dilemmas • Police Ethics • Proactive Law Enforcement • Undercover Policing

## **1 Dirty Hands and Undercover Policing: Nathan's Criticisms**

It is often said that politics is a dirty business. To be somewhat more precise: the 'dirty hands' model (henceforth, 'DHM') of politics asserts that it inevitably involves situations where whatever the politician does is morally problematic. One might think that this analysis extends to the exercise of state power in general. Policing (or, more broadly, law enforcement) seems like an obvious example, particularly when it deploys coercive, violent, covert, or undercover methods or practices that arouse moral unease. The police arguably use such methods to pursue valuable ends, like the prevention, detection, and reduction of crime. Policing thereby creates situations where perhaps whatever police officers do will have morally problematic aspects. DHM might therefore be considered a fitting theoretical framework within which to analyse the morality of some central aspects of reactive and proactive policing.

Following Christopher Nathan (2017; cf. 2022: 30–32), we focus mainly upon a type of proactive policing that arouses moral unease and to which DHM might be considered applicable, namely undercover policing. Nathan (2017: 37) introduces, but rejects as 'unsatisfying', the DHM of undercover policing:

The view often attributed to Machiavelli is that power inevitably involves doing some things that are wrongs, arising from genuine moral dilemmas. We must accept this moral residue, but we also do better not to dwell on our misdeeds. On this view, committing moral wrongs is part of the core of undercover work. The best we can do is to embrace the values we gain: in this case, the reduction of crime and the increase in security. It retains, nonetheless, a tragic element, since it is necessary that the work is performed, and those who perform it commit wrongs, thereby performing a sacrifice.[...]

A public that takes on board this view of manipulative policing will correctly feel that it puts wrongful acts at the centre of police practice. The wrongs may be justified by appeal to necessity, but unease will remain. Furthermore, one can reasonably expect that the effects of an internalisation of a dirty hands ethic by agents of a practice that is inherently secretive would be to encourage further secretiveness. A belief on the part of its agents that the practice is not wrongful is more conducive to public justification.

Nathan does not examine whether DHM *applies* to undercover policing in the first place.<sup>1</sup> Since his criticisms of the DHM of undercover policing hold only if the model applies, it is pertinent to examine whether the model does indeed apply.

We do this by focusing on a specific method of proactive undercover law-enforcement, namely legal entrapment to commit a crime (henceforth, ‘legal entrapment’).<sup>2</sup> Our first aim, then, is to assess whether DHM applies to legal entrapment. We show that there is a version of DHM that fits legal entrapment. Our second aim is to assess whether, when this account is applied to legal entrapment, Nathan’s criticisms of DHM hold good. We argue that they do not. Finally, our third aim is to show that the case of legal entrapment is not special: our results extend to policing practices more generally, including such routine practices as arrest, detention, and restraint.

In section 2, we provide a definition of legal entrapment based on our previous work. In section 3, we adapt to the case of legal entrapment work by János Kis on political morality and dirty hands. This enables us to set out three accounts that, fitting Nathan’s own depiction,

---

<sup>1</sup> DHM is merely the starting point for his discussion as a model we should steer away from. We take Nathan’s objections only to urge us to look for a better moral account of undercover policing, *if there is any*. We do not read him as offering a refutation of DHM *per se*.

<sup>2</sup> Undercover policing is often seen as a crucial policing method, and it is becoming increasingly ‘normalized’ (HMIC 2014; cf. Loftus 2019). While the morality of entrapment is often discussed from other angles, it is interesting to consider the viability of applying DHM to it. This is a natural development from Nathan’s depiction of how DHM might apply to undercover policing in general.

understand the ethics of legal entrapment in terms of a moral dilemma involving dirty hands. We argue that the first two accounts, adapted from Kis, are inapplicable to legal entrapment. The third, which is based on, but significantly different from, Kis's formulation, has, we argue, better prospects of applying. It, however, leaves no room for wrongful acts, as opposed to acts that merely have bad aspects, as part of the picture: it therefore precludes Nathan's fundamental criticism of DHM, namely his contention that on DHM a wrongful act is inevitable. In section 4, we consider a possible way of reinstating Nathan's fundamental criticism so that it appeals to badness rather than wrongfulness. Our response to this attempt enables us to contest Nathan's other criticisms of DHM, and to extend to other policing practices our main result, namely that there is a version of DHM that both applies to legal entrapment and resists Nathan's criticisms. Section 5 is a concluding summary.

## **2 Legal Entrapment to Commit a Crime: A Definition**

Cases of entrapment involve a party that intends to engage in entrapment, whom we call the 'agent', and a party that is entrapped, whom we call the 'target'. Let the terms 'party', 'agent', and 'target' encompass both individuals and groups. We draw two distinctions, which cut across each other, concerning acts of entrapment. The first concerns the status of the agent; the second concerns the act that the target performs and that the agent procures.<sup>3</sup>

*Legal entrapment* (or 'state entrapment') occurs when the agent is a law-enforcement officer, acting (lawfully or otherwise) in their official capacity as a law-enforcement officer, or when the agent is acting on behalf of a law-enforcement officer, as their deputy. When, on the other hand, neither of these is true of the agent, we have *civil entrapment* (or 'private entrapment').

---

<sup>3</sup> Our notion of procurement is technical: the agent has an intentional influence, via directly related communicative acts, on the target's will. See (*redacted*) for further discussion.

We distinguish between procured acts of criminal and of non-criminal types. An investigative journalist might entrap a politician into performing a morally compromising act that is not a crime in order that the journalist might expose the politician for having performed the act. When the act is non-criminal but is morally compromising (whether by being immoral, embarrassing, or socially frowned upon in some way), we are dealing with *moral* entrapment (using the word ‘moral’ in a wide sense). When the act is of a criminal type, we have *criminal* entrapment.

Accordingly, we classify acts of entrapment via the following two-dimensional matrix.

	A	B
1. Is the agent acting (permissibly or otherwise) in their capacity as a law-enforcement agent or their deputy?	Yes	No
2. Is the act that the agent intends the target to commit of a type that is criminal?	Yes	No

We thus get four types of entrapment:

- Type 1 = 1A + 2A = legal entrapment to commit a crime
- Type 2 = 1B + 2A = civil entrapment to commit a crime
- Type 3 = 1B + 2B = civil moral entrapment
- Type 4 = 1A + 2B = legal moral entrapment

Type 1 entrapment, now called ‘legal entrapment’ for short, is the kind that is relevant to our discussion. Elsewhere (*redacted*) we argue that an act is one of legal entrapment if and only if it meets the following five conditions:

- (i) a law-enforcement agent (or the agent's deputy), acting in an official capacity as (or as a deputy of) a law-enforcement agent, *plans* that the target perform an act;
- (ii) the act is of a type that is criminal;
- (iii) the agent *procures* the act (using solicitation, persuasion, or incitement);
- (iv) the agent intends that the act should, in principle, be *traceable* to the target either by being *detectable* (by a party other than the target) or via *testimony* (including the target's confession), that is, by *evidence that would link the target to the act*;
- (v) in procuring the act, the agent intends to be enabled, or intends that a third party be enabled, *to prosecute* (or to threaten to prosecute) the target for having performed the act.<sup>4</sup>

### 3 Moral Dilemmas, Dirty Hands, and Legal Entrapment

We now have a suitable notion of legal entrapment at hand. The next step is to connect it to a suitable account of DHM. How are we to understand DHM? To begin, it is useful to recall the following main elements of Nathan's description of DHM:<sup>5</sup>

- Moral wrongs (i.e. impermissible actions) are committed.
- Genuine moral dilemmas are involved.
- A moral residue is involved that we must accept.
- The overall picture is tragic, despite a reduction in crime and an increase in security, because a moral wrong is unavoidable: a 'sacrifice' must be made to achieve these gains.<sup>6</sup>

---

<sup>4</sup> Entrapment (without the qualifier) differs in that conditions (i), (ii), and (v) are more inclusive.

<sup>5</sup> Nathan lists an aspect of DHM that we do not list: that we, and undercover agents themselves, are encouraged not to dwell too much on their misdeeds. However, this aspect is a peculiar addition. It might have been true of Machiavelli's original formulation of DHM, but it is not widely accepted today, nor is it consistent with our depiction of agents in these situations.

<sup>6</sup> The sacrifice consists in acting contrary to a value: in other words, it constitutes a kind of moral damage.

Let us clarify some important terms here. We take ‘moral dilemma’ to refer to a choice situation in which the agent is confronted with moral demands and whatever course is taken is morally problematic. We do not equate ‘morally problematic’ with ‘morally wrong’: this is so as not to foreclose the important possibility that, even if acting with dirty hands involves facing moral dilemmas, these dilemmas are not best characterized in terms of moral wrongness. We will explain the meaning of ‘moral residue’ shortly.

The core of Nathan’s depiction of DHM is that acting with dirty hands involves a tragic moral dilemma in which moral wrongs are committed. That is, on a natural interpretation, DHM is presented as an offshoot of a tragic understanding of moral dilemmas. Following Kis (2008, Chapter 9), we explain three dilemmatic accounts of DHM.<sup>7</sup> We make some modifications, which we explain as we go along, to Kis’s discussion, and we somewhat simplify his presentation. Our modifications are substantial in the case of the third dilemmatic account of DHM. We consider how, if at all, each of these three dilemmatic accounts of DHM can be applied to the case of legal entrapment. Unlike Kis and others in the literature on dirty hands and/or moral dilemmas, we question an account only when this directly serves our primary purpose of examining the applicability or inapplicability of DHM to legal entrapment.

Of the three dilemmatic accounts, let us begin with the one that, we think, most closely matches Nathan’s depiction of DHM: *the tragic account* (TRAGIC). We begin each of the three subsections that follow with an indented summary of the account at issue.

### **3.1 The Tragic Account (TRAGIC)**

The agent, *S*, is bound by two moral demands that cannot simultaneously be satisfied.

Whichever demand *S* disregards, *S* violates a valid, in-force demand. The dilemmatic

---

<sup>7</sup> Of course, there are other ways of conceptualizing DHM: see Coady (2018), Nielsen (2000); de Wijze 2007; Nick (2022). Focusing on Kis’s discussion, however, serves our dialectical purposes well.

situation is inescapable, in that even if  $S$  becomes involved innocently  $S$  cannot come out of it innocently. However  $S$  acts,  $S$  will act impermissibly and incur guilt.

Recall that Nathan (2017: 37) says that situations involving dirty-handed acts retain a tragic element ‘since it is necessary that the work is performed, and those who perform it commit wrongs, thereby performing a sacrifice’. There are two necessity claims here. The first claim is one of *instrumental necessity*. Nathan assumes that undercover policing methods, even if they are methods of last resort, are necessary means towards their ends. While it is doubtful that the actual deployment of undercover policing typically meets this condition (cf. Loftus 2019), we assume, for the sake of argument, that the necessity towards such ends is a normative constraint on uses of undercover methods: that is, we assume that they are permissible only when instrumentally necessary towards their ends. The second form of necessity to which Nathan seems to allude is *moral necessity*: when the use of undercover methods is instrumentally necessary, the police have a moral duty to use them.<sup>8</sup> This then leads to familiar slogans that aim to bring out the paradoxical nature of the resultant situation: that sometimes it is right to do what is wrong (de Wijze 2007), or that sometimes whatever we do is wrong (Coady 2008).

Slogans, however, do not help us to understand the underlying structure of the situation. Let us, therefore, first formalize the account, closely following Kis (2008: 238). We could say that the following happens in such tragic situations. Let  $S$  be the agent (in our case, the law-enforcement officer or their deputy), let  $a$  be a course of action that involves entrapment, and let  $b$  be an option that does not. In order to generate a dilemmatic situation, an analysis must begin with the following three assumptions:

---

<sup>8</sup> Unless the second type of necessity is moral, we do not see how the DHM of undercover policing would involve, as Nathan takes it to involve, a *moral* dilemma.



A1: There is a moral demand that *S* should perform *a*, and there is a moral demand that *S* should perform *b*.

A2: *S* can satisfy each of the two demands separately.

A3: *S* cannot satisfy both demands together.<sup>9</sup>

This gives us a situation of moral conflict, but not yet a situation of a moral dilemma. For that we require a further assumption:

A4<sup>TR</sup>: Of the two demands, neither overrides the other: both emerge undefeated.

A4<sup>TR</sup> is also crucial because it partially explains the tragic (and paradoxical) nature of the situation: there is no way that the agent can do the right thing without also doing something wrong. This, however, is only a partial explanation. According to Kis (2008: 239), moral dilemmas are tragic largely because of the agent's lost innocence. The idea is that dilemmas are *inescapable*: one finds oneself in the dilemmatic situation through no fault of one's own. That is, one goes into the situation innocently, but, because of the nature of the choice involved, one cannot come out of it innocently. Given the moral demand on each side of the dilemma *and* that the dilemma is inescapable, TRAGIC has four tragic implications. To paraphrase Kis (2008: 239), these are:

I1<sup>TR</sup>: Whichever moral demand *S* chooses to disregard, *S* violates a valid, in-force moral demand.

I2<sup>TR</sup>: The dilemmatic situation is such that *S* may become involved in it innocently.

---

<sup>9</sup> Under the influence of Railton's (1996) discussion, Kis (2008: 238) characterizes moral dilemmas as involving conflicting *oughts*. We have instead characterized dilemmas in terms of conflicting *moral demands*: we take this to be closer to the spirit of Railton's discussion, and a better way of attaining generality.

I3<sup>TR</sup>: Once in the dilemmatic situation, *S* has no opportunity to come out of it innocently.

I4<sup>TR</sup>: Whether *S* performs *a* or *b*, it will be appropriate for *S* to feel guilty.

On this account, moral dilemmas involve three layers of the tragic. The first is encapsulated in I1<sup>TR</sup> (and based on A4<sup>TR</sup>), and the second in the conjunction of I2<sup>TR</sup> and I3<sup>TR</sup>. Once these are in place, I4<sup>TR</sup>, the third layer, falls into place.

The question that concerns us, in relation to each of the three accounts of DHM, is whether it applies to the case of legal entrapment (and, if so, whether Nathan's criticisms of DHM apply specifically to the case of legal entrapment). In relation to TRAGIC, the applicability of each of the three layers of the tragic can be called into question. We focus on the first two. The third layer might not be supported by the phenomenology of these cases, i.e., if law-enforcement agents that entrap typically do not feel guilty about what they have done.<sup>10</sup> Of course, this does not rule out the view that it would be *appropriate* for them to feel guilty. Since the appropriateness of such reactive moral emotions is determined by the moral structure of the case, the third layer of the tragic is dependent upon the previous two layers.

With respect to the first layer, it is far from clear that A4<sup>TR</sup> (hence I1<sup>TR</sup>) holds true for most cases of legal entrapment. Certainly, law-enforcement officers that entrap might be doing what is *normally* considered wrong overall (i.e. all-things-considered); after all, for example, they often tempt and deceive people, not always career criminals, into doing something criminal. Still, in the context of legal entrapment, it is far from clear that A4<sup>TR</sup> holds true: given all the good that might be achieved through an act of entrapment (e.g., in terms of long-term crime prevention), it might be that the moral demand not to entrap *is* sometimes overridden. While the literature on legal entrapment contains conflicting views about its permissibility, the

---

<sup>10</sup> Even if they do feel *bad* about what they did, there can be other explanations of this fact. They might doubt that they have done the right thing (cf. Nielsen 2000). Other explanations of bad feelings (regret and remorse) are possible, and we appeal to them in what follows.

literature does not suggest (at least to our knowledge) that no entrapment scenario gives rise to a clear normative conclusion about entrapping as against not entrapping. In short, even on a charitable approach, there is reason to hold that A4<sup>TR</sup> is not generally true of scenarios of legal entrapment. Rather than taking TRAGIC to apply across the board, it seems that a case-by-case analysis would be needed.<sup>11</sup>

With respect to the second layer, the inescapability requirement is of dubious applicability to cases of legal entrapment. After all, legal entrapment involves law-enforcement agents (or their deputies) choosing to entrap their targets. Although a law-enforcement agent might be ordered by a superior to entrap a target, the choice to entrap in this situation can hardly be construed as inescapable (and certainly not on the part of the superior). In addition, both the superior and the subordinate, we can reasonably assume, made, or at least could have made, an informed choice when they became law-enforcement agents. It is hard to believe that those recruited into the police could not at least foresee, upon reflection, what might await them in the service. Police officers receive training, and this training involves information on the different aspects of the job. This applies even more to those working undercover: they receive specialist training, and are recruited from the ranks of ‘ordinary’ police officers. It might be that some agents stumble into being asked to engage in entrapment in jurisdictions in which entrapment is generally discouraged; still, this does not make their choice, in any relevant sense, inescapable.<sup>12</sup> It is reasonable to conclude, then, that the decision to engage in entrapment is, at least typically, a free and informed one. By contrast, the stereotypical case of an inescapable choice, which is often used in discussing TRAGIC, is Sophie’s choice (from Styron 1979). Sophie had to choose which of her two children was to be sacrificed, and if she refused the

---

<sup>11</sup> Cf. Alexandra (2000) on typical cases of ‘noble cause’ corruption. Walzer (1973) is unclear on exactly what situations are covered, but his focus is mostly on emergencies, which, by their nature, are rare situations.

<sup>12</sup> One could say that policing *needs* to be done—a *social* choice—so *someone* must dirty their hands. While this may be true in politics, it is far from clear that we *need* entrapment or undercover policing in general. Besides, there is something morally problematic about making an *individual* choice inescapable on the basis of a (presumed) *social* choice.

choice both children were to be taken to the gas chamber. The choice, moreover, was imposed on her by Dr Mengele: Sophie neither created the choice situation, nor intended it, nor had foreseen it, nor had an innocent way out of it. Sophie, thus, truly loses her innocence no matter what she does. It is doubtful, to say the least, that we can say anything remotely similar of law-enforcement officers that engage in entrapment.

Lastly, Nathan associates a *moral residue* with the ‘dirty hands’ situation. Now, this is true in one way and wrong in another. If the agent has done wrong, then the agent, per I4<sup>TR</sup>, is guilty of wrongdoing, and this does not seem to us to constitute a moral remainder. True, in the literature on ‘dirty hands’ many would be fine with this. They would just claim that ‘moral residue’ here is referring to the moral phenomenology of these cases, i.e., to what the agent experiences, or, at least, to what it would be appropriate to experience. We are, however, critical of this approach. We do not consider it plausible to construe the moral residue as merely phenomenal. Instead, we think, the relevant residue should be located in the moral structure of entrapment, and not (merely) in its phenomenology. More generally, we suggest, it is the moral structure of a scenario that determines what it is appropriate for a person in that scenario to feel. Nevertheless, Nathan’s mention of moral residue suggests an alternative account that might be considered applicable to legal entrapment: the *moral residue account* (RESIDUE).

### **3.2 The Moral Residue Account (RESIDUE)**

*S* is bound by two moral demands, to perform *a* and to perform *b*, that cannot simultaneously be satisfied. The demand to perform *a* overrides the demand to perform *b*, but *b*’s normative force does not evaporate: rather, it gives rise to a derivative requirement that the target of *S*’s act must receive redress.

Williams (1965) has argued that what happens in moral dilemmas is that, contrary to A4<sup>TR</sup>, one moral demand overrides the others, but the overridden demands are not silenced. They ‘stick around’; their force does not evaporate.<sup>13</sup> In particular, these defeated demands generate *derivative demands* to compensate for, or to repair, the damage done: this is the moral residue that TRAGIC lacks. As Kis (2008: 251) puts it:

The defeated ought has no action-guiding force in the immediate context of the situation in which the choice is being made, but it has action-guiding force in the context of a later choice that emerges in virtue of *S*’s action.

That is, the decision situation is more complex than in TRAGIC. If agent *S* opts to perform *b*, then *S* must perform *c<sub>a</sub>* (where this represents compensating for, or repairing the harm caused by, having failed to perform *a*). If, on the other hand, *S* opts to perform *a*, then *S* must then perform *c<sub>b</sub>* (where this represents compensating for, or repairing the harm caused by, having failed to perform *b*). The decision is more complex because, when deciding how to act, *S* must not only decide whether to do *a* or *b*, but also which, if either, of *c<sub>a</sub>* and *c<sub>b</sub>* is a feasible option.

Before we address the question of application, let us this time highlight an important general problem: it appears that RESIDUE’s claim to a moral residue vanishes on closer analysis (Kis 2008: 252). After all, if *S* makes the right choice by choosing to act on the overriding demand and also compensates the victims of this choice (for the failure to take the other courses of action that were initially open), then no moral residue remains. *S* simply did the right thing, on both levels (acting and then compensating): the moral universe remains intact, and *S* comes out of the situation (morally) innocent.

---

<sup>13</sup> RESIDUE is supported naturally by a picture of competing *pro tanto* reasons the balancing of which gives us an all-things-considered ought-judgement. See Alexandra (2000), who depicts ordinary cases of ‘noble cause’ corruption in exactly this way.

While there is a way around this problem, it suggests that the proper form of RESIDUE whose application to legal entrapment we intend to query is not the one with which we began. What is the way around the problem? As Kis (2008: 253) points out, if there is *irreparable* damage involved in a choice situation, then a moral residue would *necessarily* remain. (Kis calls this ‘non-eliminable moral residue’.) This appears to restore the tragic character of *S*’s choice-situation, since there would be no way for *S* fully to satisfy the requirements that apply. To formalize it, the following account of the ‘moral residue’ account would hold. Let us start with the original account of moral conflict:

A1: There is a moral demand that *S* should perform *a*, and there is a moral demand that *S* should perform *b*.

A2: *S* can satisfy each of the two demands separately.

A3: *S* cannot satisfy both demands together.

We require, following this version of RESIDUE, a fourth assumption (Kis 2008: 253):

A4<sup>MR</sup>: Performing *a* involves a non-eliminable moral residue, and either performing *b* involves a non-eliminable moral residue or the demand to perform it is not overriding.

Replacing A4<sup>TR</sup> with A4<sup>MR</sup> produces its own problems, however. First, as Kis (2008: 253) points out, if damage is irreparable, i.e., if it *cannot* be repaired, then, if ‘ought’ implies ‘can’, it is not the case that it ought to be repaired. This means that the ‘rediscovered’ tragic element in RESIDUE becomes diluted. If, on the one hand, the damage is reparable, then no residue need remain. If, on the other hand, the damage is irreparable, then a residue remains but redress

is not required (because it is impossible). In either case, no derivative moral requirement remains that could, if violated, trigger a tragic dénouement.<sup>14</sup>

Secondly, Kis (2008: 253) argues that this does not rule out the possibility that it would be appropriate for *S* to feel bad about having failed to compensate the victim(s) of *S*'s act. This feeling should not be guilt, and perhaps not even regret or remorse. Still, *S* can think of the act as morally problematic and feel bad about this. This gives a thinly tragic analysis: whatever *S* does, it is appropriate for *S* to feel bad about the chosen act. But, first, it could be asked exactly what Kis has in mind here: what would 'feeling bad' amount to and why would it be appropriate? And, secondly, as far as the moral analysis of entrapment is concerned, this leaves us with very little of the tragic aspect of the situation as Nathan originally described it.

Furthermore, thirdly, while the original version of RESIDUE was generally applicable to legal entrapment, the present version applies only to those instances of it that involve irreparable damage. Arguably, however, most cases of legal entrapment are too mundane to involve irreparable damage: law enforcement does not normally involve killing or maiming entrapped subjects, for example. In fact, this gets worse if we consider that, on A4<sup>MR</sup>, the choice not to entrap should be either not overriding (which is not obvious) or should cause irreparable damage (which is not generally the case). In short, what we gain in 'tragedy' by changing focus to irreparable damage, we lose in scope of application.

Lastly, there is good reason to think that RESIDUE fails to preserve the dilemmatic nature of dirty hands. As Kis (2008: 255–256) shows, reference to irreparable damage cannot be what constitutes moral dilemmas because hard choices of the form depicted by A1–A3 that are *not* moral dilemmas *can also* involve irreparable damage. For example, a rescuer might

---

<sup>14</sup> One could claim that the reason why we need to compensate victims of wrongdoing, even if the damage is irreparable, is communicative: e.g., to signal to both victims and the wider society that the wrongdoer acknowledges the importance of the moral value that was violated. This might also be behind existing court practices to provide entrapment remedies (*redacted*). The question, however, is whether such pragmatic, non-moral grounding would be enough to reinstate the tragic nature of RESIDUE. This seems far from obvious to us.

allow someone to die by deciding to save another and the choice might be perfectly well supported by moral reasons. RESIDUE fails to distinguish such cases from genuine moral dilemmas.

RESIDUE, we conclude, is not a viable dilemmatic account of DHM suitable for application to the case of legal entrapment. Let us turn, then, to the third account drawn from Kis's discussion—the '*dirty hands*' account (DIRTY).

### 3.3 The '*Dirty Hands*' Account (DIRTY)

*S* is bound by two moral demands, to perform *a* and to perform *b*, that cannot simultaneously be satisfied. Although the demand to perform *a* overrides the demand to perform *b*, performing *a* remains morally bad. Hence, while it is right for *S* to choose to perform *a*, what *S* does remains morally bad.

The above summarizes our version of this account, which differs significantly from Kis's. We now explain Kis's version, and then, using it as background, explain and justify the details of our version. We start with the usual three assumptions to depict moral conflict:

A1: There is a moral demand that *S* should perform *a*, and there is a moral demand that *S* should perform *b*.

A2: *S* can satisfy each of the two demands separately.

A3: *S* cannot satisfy both demands together.

The fourth assumption, as before, is the one that fulfils the task of accounting for the dilemmatic nature of the situation including dirty-handed acts (Kis 2008: 264). *Moral reprehensibility* is, on Kis's account, the central concept that features in this assumption (Kis 2008: 267):



A4<sup>DH</sup>: Performing *a* is morally reprehensible, and performing *b* is either morally reprehensible or the demand to perform it is not overriding.

Kis (2008: 260–263) explains how A4<sup>DH</sup> is meant to help secure a dilemmatic account of DHM. This involves five central ideas. The first concerns the aforementioned notion of *moral reprehensibility*. Kis holds that an act can be right, hence morally acceptable, in certain circumstances, and nonetheless morally reprehensible in the same circumstances. This is possible, says Kis introducing his second central idea, because some acts, such as murder and betrayal, have *essential properties* that make them morally reprehensible irrespective of the circumstances.<sup>15</sup> Some concepts, like *murder* or *betrayal*, Kis says, have descriptive content that cannot be pried apart from evaluative criteria: the acts they describe cannot be identified in morally neutral terms. Murder and betrayal, Kis asserts, remain morally reprehensible even if in the given situation they are, all things considered, the morally right thing to do.

The third idea is that *threshold deontology* is an appropriate approach to normative ethics (cf. Nagel (1986); Coady (2018: §7); Alexander & Moore (2021: §4)). At its base, the idea is that we can evaluate an act in two ways: according to the states of affairs it produces (the viewpoint typically associated with consequentialism) and according to how it treats its object (the viewpoint typically associated with deontology). It is the latter that is crucial: if an act fails to treat its object as it should be treated then this makes it morally reprehensible. Now, action-based constraints, on the ‘threshold morality’ view, often outweigh concerns related to consequences: the putatively good consequences that the action would bring about are outweighed by the constraints relating to how the action would treat its object. Nevertheless, beyond a certain threshold (e.g. avoidance of great harm) considerations of the consequences override those that relate to the intrinsic nature of the action. Importantly, even in such cases,

---

<sup>15</sup> This idea is based on what we think is a debatable reading of Marcus (1996).

the concerns pertaining to the nature of the action remain in place as *evaluative* considerations: it is morally right to avoid great harm, but what is, with respect to the nature of the action, inappropriate treatment, nevertheless remains inappropriate treatment.<sup>16</sup> Hence the act, albeit morally right because justified by its consequences, remains morally reprehensible: this is a moral residue. Kis's fourth central idea is that this makes DIRTY paradoxical: whatever the agent does will be morally acceptable and unacceptable at the same time.

Finally, we come to Kis's fifth idea: although the acts the reprehensibility of whose natures is overridden by the goodness of their consequences in dilemmatic situations are morally acceptable (because right) and unacceptable (because reprehensible) at the same time, they are not blameless and blameworthy at the same time. The justified nature of the act means that no blame is appropriate. The act is morally reprehensible, but the proper response to this is not blame, but *regret* or *remorse*. An observer's proper responses are not resentment and indignation, but fear and pity. Kis (2008: 265) depicts the appropriate phenomenology of dirty-handed acts as lying between cases of faultless involuntary contributions to accidents (where only what Williams (1976) calls 'agent-regret' is appropriate) and blameworthy wrongdoing (where guilt and blame are appropriate).

While Kis's formulation of DIRTY would, we think, apply to legal entrapment, our interest is in a version of DIRTY that expunges some commitments from Kis's formulation. Getting rid of these commitments has, we think, two main virtues: first, it renders the amended account more plausible; secondly, it expands the number, and the variety, of morally interesting situations to which DIRTY can be applied.<sup>17</sup>

---

<sup>16</sup> It is not taken up in Kis's presentation exactly how threshold deontology sees the inter-relation of reasons grounded in the consequences and reasons grounded in the intrinsic nature of the action. Raz (1999)'s notion of exclusionary reasons seems to be in the background here, though it is notoriously hard to interpret (Adams 2021).

<sup>17</sup> In a revised edition of the book, available only in Hungarian, Kis (2017: 290–2) discards threshold deontology in favour of a distinction borrowed from Scanlon (2008: Chapter 1) between the 'critical question' (regarding the agent's attitudes) and the 'deliberative question' (regarding the agent's acts). Kis does this because he wants to avoid the paradoxical nature of his earlier position. Our discussion is not influenced by this change to Kis's position. Our formulation of DIRTY retains threshold-deontology while avoiding both inconsistency and the idea of moral reprehensibility.

We amend Kis's first commitment so that the notion of reprehensibility is replaced with the idea that an act can have a *morally bad aspect*. Thus, on our version of DIRTY, an act can be right, hence morally acceptable, in certain circumstances, but nonetheless have a morally bad aspect that, in the same circumstances, does not evaporate. Our next move is to specify further the nature of such aspects. Here, too, we disagree with Kis, now concerning his second central idea: essential properties. Besides holding that serious critical questions can be asked about Kis's understanding of essentialism,<sup>18</sup> we also reckon that if a viewpoint in normative ethics can forswear a metaphysical commitment, while still doing its normative-ethical work, then it is better for it to do so. In any case, we think that it is not the idea of an act's having essential properties that is, at bottom, relevant to deontology. Instead, it is the familiar idea that, apart from having *extrinsic* consequences (which they bring about), some acts have *in themselves*, as tokens of their types, morally relevant *intrinsic* features (such as being a promise, the breaking of a promise, or a lie). An appropriate example might be if a field surgeon, to a patient's great extrinsic benefit, in order to prevent a tragic outcome for the patient, and under orders, violated the patient's autonomy by administering a treatment to which the patient had expressly not consented. This act would have, we surmise, overwhelming extrinsically good consequences (in that, so to speak, it would save life or limb), while it would also be intrinsically morally bad (in that it would violate the patient's autonomy).

It is this notion of *intrinsic moral badness* that we propose to use instead of Kis's notion of *essential moral reprehensibility*. We are now in a position to replace Kis's A4<sup>DH</sup> with our own proposal:

---

<sup>18</sup> A venerable approach to metaphysics regards *substances* as things of the right kind to have essences; acts would not count, according to this tradition, as things of the right kind. Furthermore, it is clear that Kis is working with a *modal* account of essence, according to which all and only those properties that an entity has 'in all possible worlds', or as a matter of necessity, are of its essence. That conception of essence has often been contested, especially since Fine (1994).

A4<sup>DH\*</sup>: Performing *a* has a morally bad aspect, and performing *b* either has a morally bad aspect or the demand to perform it is not overriding.

We next deploy this notion within Kis's third idea: threshold deontology. The example above of the field surgeon is a case at hand: given the overwhelming positive consequences (even when balancing them against the negative consequence of pain), the threshold is reached, and administering the treatment becomes the morally right thing to do, despite the fact that doing so is intrinsically morally bad because it violates the patient's autonomy (since the soldier has not consented to the treatment).

Furthermore, given A4<sup>DH\*</sup>, we diverge from Kis's view that an act can be both morally acceptable (because right) and morally unacceptable (because reprehensible) at the same time.<sup>19</sup> We supplant Kis's notion of moral reprehensibility, and his subsequent paradoxical deployment of the notion of moral unacceptability, with the notion that, in the dilemmatic situation, a threshold-deontological account sees one course of action as *right but intrinsically bad*. In the situations to which our amended version of DIRTY is meant to apply, the following happens. When the agent acts *rightly*, intrinsic moral badness is nevertheless produced. When the agent acts *wrongly*, an intrinsic moral good is promoted, but a more significant extrinsic ill is brought about. In short, on DIRTY, while the agent can choose to act rightly, they cannot avoid doing something morally bad at the same time. This left-over intrinsic moral badness is a moral residue that acting rightly leaves behind.

Finally, turning to Kis's fifth central idea, from a phenomenological point of view, how should the agent then regard their action? Here we think that Kis's phenomenological picture can be weakened somewhat. We agree that in DIRTY it *can be* appropriate for the agent to feel

---

<sup>19</sup> This paradox, or seeming paradox, might be thought to be an undesirable feature of Kis's account. This perhaps raises bigger questions, which we do not propose to address, about the appropriate philosophical methodology to use when approaching a seeming paradox.

remorse at having to take the course of action in question, even though taking that course is permissible (and perhaps even mandatory) in the circumstances, and that this remorse is something more than mere agent-regret (Williams 1976), though less than guilt (perhaps it is tragic-remorse, see de Wijze's 2004). On our account, however, remorse would be appropriate only in situations in which the agent's chosen course involved a grave moral ill. Similarly, observers in such serious cases can react with fear and pity as Kis proposes. In other, less serious, situations, perhaps moral disappointment would be appropriate (cf. Menges 2020).

For our modified version of DIRTY to apply to legal entrapment, the act of entrapment must be intrinsically morally bad in the sense explained above. One candidate is Howard (2016: 25) who argues that entrapment (whether legal or civil):

*subverts* the moral capacities of entrapped persons. To subvert an agent's moral capacities is to interfere with the agent's practical reasoning in ways that increase the likelihood she will culpably choose to act wrongly. Such activity [...] is incompatible with respect for that agent. Specifically, it is incompatible with [...] an attitude of support for the successful operation of others' moral capacities.

While we are inclined to think that there are other morally bad aspects of legal entrapment, some of which may be agent-centred rather than target-centred,<sup>20</sup> we agree with Howard that such subversion is involved in legal entrapment, and that it is morally undesirable.<sup>21</sup> While Howard sees the subversion that is involved as sufficient to make entrapment *wrong*, rather

---

<sup>20</sup> On this distinction, and on some agent-centred objections to entrapment, see (*redacted*). For details of further objections, see Dillof (2004), Hughes (2004), Carlon (2007).

<sup>21</sup> It is crucial for Howard that the target 'will culpably choose to act wrongly'. Like some courts elsewhere (see *redacted*), Howard is therefore implicitly rejecting the specific entrapment doctrine that has become orthodox in the US. We also reject that doctrine. Moreover, in saying that entrapment subverts the target's moral capacities short of nullifying culpability, Howard would be able to avail himself of the distinction (explained in *redacted*) between *causing* the target to commit an act (e.g. by the administration of a mind-altering drug) and *procuring* the act (as we there understand it).

than just *bad*, within the version of DIRTY that we are entertaining it is seen merely as a bad aspect. When is this badness outweighed, rendering legal entrapment permissible? As in Kis's discussion, this occurs when great harm is at stake. Now, there are many mundane cases where the target of legal entrapment is unlikely to have otherwise proceeded to cause great harm; legal entrapment is unlikely to be justified in such cases. On the other hand, the level of subversion inherent in a particular act of legal entrapment might also be relevant: if this is minor, then entrapment may be morally justified even for minor benefits.<sup>22</sup>

While the extent of DIRTY's applicability as justification for real-life cases of legal entrapment is thus unclear, this does not restrict its theoretical applicability: only those cases of entrapment are permissible where the intrinsic moral badness of the action is outweighed by the (overall) goodness of the consequences (and hence one in-force moral demand overrides another). Even in these cases, the account is non-consequentialist. This is because it remains the case that the action of legal entrapment has intrinsic negative moral value that the gains *do not nullify*. DIRTY not only explains why legal entrapment is sometimes permissible, but also why it is aptly regarded (by the courts, e.g.) as a method of last resort. Legal entrapment is still bad, and the fitting response to having done it is, under at least some circumstances, remorse.

No doubt, many theoretical questions can be asked about DIRTY, but, as before, we have tried to stay clear of the general debate concerning moral dilemmas and dirty hands. Our interest has been primarily in the application of DIRTY to legal entrapment. Unlike TRAGIC and RESIDUE, DIRTY appears to be a good candidate for this. It has no place for the idea that whatever the agent in a 'dirty hands' scenario does will be morally wrong. It is therefore immune to Nathan's fundamental criticism of DHM. At the same time, it retains not only the dilemmatic nature of the choice situation, but some element of the tragic (when the intrinsic moral badness qualifies as a grave moral ill), and a moral remainder as well.

---

<sup>22</sup> This connects to the discussion between those that advocate for a fixed threshold and those that argue for a sliding-scale form of threshold deontology. See Alexander & Moore (2021: §4).

### 3.4 Summary

Our substantive discussion took as its starting point Nathan's criticisms of the DHM of undercover policing (Nathan 2017: 37). Nathan made four assertions about what would happen if the DHM of undercover policing were to be publicly endorsed, which we now enumerate:

- 1) The public would correctly feel that morally wrongful acts were at the centre of police practice.
- 2) Despite the justification of these acts, public unease would remain.
- 3) The police would become, because of internalization of this ethic, even more secretive.
- 4) By contrast, if there were a better model for understanding the morality of undercover policing that would yield the belief on the part of its practitioners (the police) that it were not wrongful, that would be more conducive to public justification.

2) and 3) are social-scientific predictions. 1) is too, but it incorporates a theoretical assertion that we have been subjecting to philosophical scrutiny: namely, that DHM takes wrongful acts to be at the centre of police practice. This is Nathan's fundamental criticism of DHM.

We have highlighted Nathan's characterization of DHM (specifically, of undercover policing) as morally dilemmatic, tragic, and involving a moral residue. We have examined whether there was a way to flesh out Nathan's remarks into a fully-fledged account that was applicable, in particular, to legal entrapment. In doing so, we have deployed three accounts of the DHM of politics that we have adapted from Kis (2008). We have argued that the first two, TRAGIC and RESIDUE, are inapplicable to the case of legal entrapment. We have proposed a modified version of the third account, DIRTY, and we have argued that it is applicable to legal entrapment. Crucially, however, DIRTY, whether in Kis's original form or in our modified version, rejects the idea that whatever the agent in a 'dirty hands' scenario does will be morally

wrong. Nathan's case against the DHM of undercover policing was predicated upon the contention that, on DHM, undercover policing involved moral wrongdoing. Of the three accounts here surveyed, then, DIRTY remains a good choice in the face of Nathan's criticism.

#### **4 Policing, Dirty Hands and Public Justification**

At this point, a possible rejoinder is that Nathan's fundamental criticism can be revised by replacing the appeal to *moral wrongness* (from his characterization of DHM) with an appeal (from our version of DIRTY) to what is intrinsically *morally bad*. Accordingly, DHM would be characterized via the following list (with changes italicized):

- Acts with *morally bad aspects* are committed.
- Genuine moral dilemmas are involved.
- A moral residue is involved that we must accept.
- The overall picture is tragic, despite a reduction in crime and an increase in security, since a *moral evil is unavoidable*: a 'sacrifice' must be made to achieve these gains.

The amended criticism could then be summarized as follows (only 1) and 4) changes):

- 1\*) The public would correctly feel that *acts with morally bad aspects* were at the centre of police practice.
- 2) Despite the justification of these acts, public unease would remain.
- 3) The police would become, because of the internalization of this ethic, even more secretive.
- 4\*) By contrast, if there were a better model for understanding the morality of undercover policing that would yield the belief on the part of its practitioners (the police) that it *did*



*not involve acts with morally bad aspects*, that would be more conducive to public justification.

We think, however, that neither 1\*) nor 4\*) is a serious problem, and, partly for the reasons that support this suggestion, neither 2) nor 3) grounds a serious objection.

Let us start with 1\*). The first thing to notice here is that DHM, under our modified version of DIRTY, applies, at least in principle, not only to the comparatively exotic, and certainly specialized, domain of undercover or covert proactive policing, but also to any forceful or coercive elements of policing that are necessary to the practice of law-enforcement. Among the many such forceful and coercive elements of policing, consider such routine reactive practices as arrest, detention, and restraint. Although they may be for the greater good of their targets, and/or of society, these practices, given their forceful or coercive natures, can justifiably be resorted to only when the probability is relatively low that their ends can effectively and efficiently be achieved by non-forceful and non-coercive means. This is because when an act is forceful or coercive, that is an intrinsically morally bad aspect of it.

Consequently, unlike in the case of the original version of Nathan's criticism, which was predicated on the idea that morally wrong acts were central in DHM, the contention that acts with morally bad aspects are at the centre of police practice is neither specific to the case of undercover policing nor, we suggest, particularly contentious. Indeed, the observation that even everyday police practice involves morally bad aspects is innocuous, realistic, and compatible with conscientious, but morally careful, policing. To deny it would seem naïve. Moreover, the contention that policing involves acts with morally bad aspects is not something that is distinctive of DHM: it arguably follows from any realistic picture of the ethics of everyday policing. In short, our modified version of DIRTY provides an account of DHM that applies not only to legal entrapment, but to every method or practice of policing, whether

undercover or not, that involves intrinsically bad aspects. These considerations suggest that 1\*) is true but innocuous: it is ineffectual as a criticism of the DHM of undercover policing.

Turn now to 4\*). The relationship between dirty-handed acts and public justification has two sides. On the one hand, there is the side, emphasized by Nathan, of the agent and the institution of which they are a part. The agent gets dirty hands; thereby, so does their institution. The resulting demand for public justification puts a heavy burden on the agent's institution (in this case, the police). Now, if *wrongful acts* are crucial to undercover policing, then of course this makes public justification of it difficult. Nathan is right about this. On our version of DIRTY, however, the relevant acts are not morally *wrong*, but only morally *bad*. This makes public justification easier, and perhaps much easier, than Nathan envisages. In fact, given the above-noted wide scope and generality of DIRTY, it is difficult to see what one could propose as a realistic and non-naïve contrasting alternative.

It is, however, the other side of the relationship that is even more important for the evaluation of 4\*): the side of the public to whom justification is owed. If policing did not centrally involve acts with morally bad—or, for that matter, morally wrong—aspects, then it would not, as against more typical forms of labour, be in special need of moral and political (as against economic) public justification. That is, in an important sense, 4\*) gets things backwards, and Nathan's objection can be inverted. He emphasizes the *consequences* of embracing the 'dirty hands' ethic for public morale, as well as for police morale. The very ethic he targets, however, has resources<sup>23</sup> in it to control these detrimental effects, and these resources centre exactly on the notion that Nathan finds detrimentally affected: public justification. That is, *exactly because* the acts involved are dirty-handed, public accountability is placed centre stage in DHM.

---

<sup>23</sup> Liberal democracy with a strong independent media is a good example. In the case of the police, further institutional measures include all the ways of overseeing police work, the extensive discretionary rights of police officers, their original authority coming directly from the law, and other procedural barriers on police work.

It is not difficult to understand why this is so. If an agent commits a morally bad—let alone morally wrong—act, then the agent owes, at a minimum, an explanation-cum-justification to the targets (or victims) of their act. Now, from this it does not automatically follow that this must take a *public* form.<sup>24</sup> Still, there are good reasons to think that it should do so (cf. Kis 2008: Chapter 8). First, there is the problem of *moral corruption*. On the one hand, in policing we need—if we go along with the idea that there are morally justified dirty-handed acts—people that are willing to dirty their hands. On the other hand, we do not want these people to dirty their hands *too easily*. Lord Acton’s apothegm ‘power tends to corrupt, and absolute power corrupts absolutely’ is also true in policing (cf. Miller 2016). This is further underlined by a second reason: *uncertainty*. No police officer can be sure that when they entrap, for example, their reasons are indeed good. Hence, just as with moral corruption, we do not want to make it *too easy* for police officers to act on their reasons, however good they take them to be. The two considerations also connect: those that are more easily inclined to dirty their hands are also more likely not to care that they might be acting wrongfully. In short, public justification, for good reasons, is an essential part of DHM, and not an external constraint or demand on it.<sup>25</sup>

Two further conclusions concern how the modified version of DIRTY undermines 2) and 3). In relation to 2), to the extent that the public is at ease with the idea that policing is necessary at all, the public will be, or at least ought rationally to be, at ease with the idea that the police will engage in forceful or coercive acts: for policing is, as the term, both broader (in that, for example, it includes the activities of the intelligence services) and narrower (in that it excludes, for example, educational and outreach work undertaken by the police), ‘law-

---

<sup>24</sup> Public justification has not always been the norm. The prince of Machiavelli (1998) has no inner life, for example. The politician of Weber (1994) suffers internally only. Although the ‘Catholic’ model of Walzer (1973) proposes social expression, this is primarily for the repentance of sins in order to achieve salvation.

<sup>25</sup> Tillyris (2016) argues that politicians should *not* reveal their dirt to the public. But what holds for a career politician may not hold for a law-enforcement officer. Besides, it is possible that if Tillyris is right, we have a sort of tug of war on our hands: *we* should bring the politician to account for their actions, but perhaps the politician should then try to avoid us and deceive us in response.

enforcement' suggests, an inherently forceful and coercive endeavour. Contrary to 3), the modified DIRTY account makes increased secrecy by the police *less likely* than would a contrary account that denied that morally bad acts were an inevitable aspect of policing. For, given that such acts *are* inevitable aspects of policing, and that the acknowledgement of this is crucial to the need for public (moral and political) justification of the activities of police forces, the rational and strategically appropriate attitude to those acts, on the part of the police, is to admit their presence and to justify them. Secrecy, on the contrary, would be both irrational and counter-productive in respect of public justification.

## **5 Conclusion**

We have focused on DHM as a possible framework for analysing the morality of undercover policing, specifically in relation to legal entrapment. We took as our starting point the criticisms of the model by Nathan (2017). We had two main aims. Our primary aim was to see if the model applied to legal entrapment; our secondary aim was to establish whether, if the model applied, Nathan's criticisms of it held. With respect to the first aim, we presented three possible versions of the model, loosely taking our inspiration from Nathan's remarks and making extensive use of Kis (2008). We found that the first two accounts, TRAGIC and RESIDUE, were inapplicable to the case of legal entrapment. We argued that the third, DIRTY, applied, but had no room for morally wrong acts. When Nathan's criticism was amended, we found that DIRTY was immune to all four of Nathan's criticisms. In short, once we consider the range of resources available to DHM, Nathan's critique is either precluded or loses its force.

## **Acknowledgements [...]**

## References

- Adams, N.P. 2021. In defence of exclusionary reasons. *Philosophical Studies* 178: 235–253.
- Alexander, L. and Moore, M. 2021. Deontological ethics. *Stanford Encyclopaedia of Philosophy* Winter 2021 edition, ed. E.N. Zalta, URL = <https://plato.stanford.edu/archives/win2021/entries/ethics-deontological/>
- Alexandra, A. 2000. Dirty Harry and dirty hands. In *Violence and police culture*, ed. T. Coady, S. James, S. Miller and M. O’Keefe, 235–248. Melbourne: Melbourne University Press.
- Carlson, A. 2007. Entrapment, punishment, and the sadistic state. *Virginia Law Review* 93: 1081–1134.
- 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. *Stanford Encyclopaedia of Philosophy* Fall 2018 edition, ed. E.N. Zalta, URL = <https://plato.stanford.edu/archives/fall2018/entries/dirty-hands/>
- Cooper, J.A. 2012. Noble cause corruption as the consequence of role conflict in the police organisation. *Policing and Society* 22: 169–184.
- de Wijze, S. 2004. Tragic remorse—the anguish of dirty hands. *Ethical Theory and Moral Practice* 7: 453–471.
- de Wijze, S. 2007. Dirty hands: Doing wrong to do right. In *Politics and Morality*, ed. I. Primoratz, 3–20. London: Palgrave Macmillan.
- Dillof, A.M. 2004. Unraveling unlawful entrapment. *Journal of Criminal Law and Criminology* 94: 827–896.
- Fine, K. 1994. Essence and modality. *Philosophical Perspectives* 8: 1–16.

- HMIC. 2014. See also *An inspection of undercover policing in England and Wales* (HMIC 2014) <<http://www.justiceinspectrates.gov.uk/hmicfrs/wp-content/uploads/an-inspection-of-undercover-policing-in-england-and-wales.pdf>> accessed 6 June 2020
- Howard, J.W. 2016. Moral subversion and structural entrapment. *Journal of Political Philosophy* 24: 24–46.
- Hughes, P.M. 2004. What is wrong with entrapment? *Southern Journal of Philosophy* 42: 45–60.
- Kis, J. 2008. *Politics as a moral problem*. Budapest, Hungary: CEU Press.
- Kis, J. 2017. *A politika mint erkölcsi probléma*. Bratislava, Slovakia: Kalligram.
- Loftus, B. 2019. Normalizing covert surveillance: the subterranean world of policing. *British Journal of Sociology* 70: 2070–2091.
- Machiavelli, N. 1988. *The prince*, ed. Q. Skinner and R. Price. Cambridge: CUP.
- Marcus, R.B. 1996. More about moral dilemmas. In *Moral dilemmas and moral theory*, ed. H.E. Mason, 23–36. New York: OUP.
- Menges, L. 2020. Blame it on disappointment. *Public Affairs Quarterly* 34: 169–184.
- Miller, S. 2016. Noble cause corruption in policing. In: S. Miller, *Corruption and anti-corruption in policing—philosophical and ethical issues*, 39–51. Cham, Switzerland: Springer Nature.
- Nagel, T. 1986. *The view from nowhere*. New York: OUP.
- Nathan, C. 2017. Liability to deception and manipulation: the ethics of undercover policing. *Journal of Applied Philosophy* 34: 370–388.
- Nathan, C. 2022. *The ethics of undercover policing*. Abingdon: Routledge.
- Nick, C. 2022. Dirty hands and moral conflict—Lessons from the philosophy of evil. *Philosophia* 50: 183–200.

- Nielsen, K. 2000. There is no dilemma of dirty hands. In *Cruelty and deception: the controversy over dirty hands in politics*, ed. R. Rynard and D. Shugarman, 139–155. Peterborough, Ontario: Broadview Press.
- Railton, P. 1996. The diversity of moral dilemma. In *Moral dilemmas and moral theory*, ed. H.E. Mason, 140–166. New York: OUP.
- Raz, J. 1999. *Practical reasons and norms*. 2<sup>nd</sup> ed., Princeton: Princeton University Press.
- Scanlon, T.M. 2008. *Moral dimensions: Permissibility, meaning, blame*. Cambridge, Mass.: Belknap Press.
- Styron, W. 1979. *Sophie's choice: a novel*. New York, N.Y.: Random House.
- Tillyris, D. 2016. After the standard dirty hands thesis: Towards a dynamic account of dirty hands in politics. *Ethical Theory and Moral Practice* 19: 161–175.
- Walzer, M. 1973. Political action: the problem of dirty hands. *Philosophy and Public Affairs* 2: 160–180.
- Weber, M. 1994. The profession and vocation of politics. In: M. Weber, *Political writings*, tr. R. Speirs, ed. P. Lassman, 309–369. Cambridge: CUP.
- Williams, B. 1965. Ethical consistency. *Proceedings of the Aristotelian Society, Supplementary Volumes* 39: 103–124.
- Williams, B. 1976. Moral luck. *Proceedings of the Aristotelian Society, Supplementary Volumes* 50: 115–135.
- Williams, B. 1985. *Ethics and the limits of philosophy*. Cambridge, Mass.: Harvard University Press.