

# The Manifestation Account of Evil

*Philipp Schwind*

Center for Ethics, University of Zurich, Zürich, Switzerland

*philipp.schwind@philos.uzh.ch*

*Felix Timmermann*

Center for Ethics, University of Zurich, Zürich, Switzerland

*felix.timmermann@philos.uzh.ch*

## Abstract

This article defends a novel definition of evil. An action is evil if (1) a pro-attitude (or complete indifference) towards severe harm to a sentient being is (2) manifested in the action. The manifestation can take either of two forms: expressing the pro-attitude or attempting to realize its object. In order to exclude cases where the pro-attitude is the result of a positive attitude and the action does therefore not count as evil, the pro-attitude (3) must be generated from a morally reprehensible attitude such as greed or sadistic pleasure. As an implication of this definition, not every evil action is extremely bad, and some ‘merely’ wrong acts might be worse than some evil acts.

## Keywords

evil – pro-attitude – moral horror – silencing – axiology

## 1 Introduction

There are acts that display such atrocity, viciousness or cruelty that any normal observer recoils in horror. To mark these acts and to separate them from others that are ‘merely’ wrong, we call them “evil”. Despite the fact that the notion of evil action has its place in our moral vocabulary, it has proven difficult to give an adequate secular account of the term. A comparison of different actions which all deserve to be called evil demonstrates why this is so. Consider the following three cases:

- (1) While working as a nurse in Oldenburg and Delmenhorst in northern Germany from 1999–2002 and 2003–2005, Niels Högel administered unauthorized injections that triggered cardiac arrest on numerous patients; over the years, possibly more than 100 of them died as a consequence of Högel's actions (the exact number can only be estimated). As a psychologist who evaluated him explained during court proceedings, killing for its own sake was not Högel's aim. Rather, Högel admitted that he had acted out of boredom and the desire to show off his resuscitation skills when he succeeded in bringing the patient back to life at the last moment (which earned him among colleagues the nickname "Rescue Rambo"). His choice of victims appears to have been random, with their ages ranging from 34 to 94.<sup>1</sup>
- (2) In an attempt to extort 11.75 million Euros, Jochen S. placed poisoned jars of baby food on shelves in five supermarkets in the city of Friedrichshafen and told retailers he would identify the items in exchange for money. He also threatened to place 20 more poisoned food jars. The jars contained *ethylene glycol*, an odourless and colourless toxic liquid used in antifreeze for cars; its sweet taste can be attractive to children. If swallowed, the substance leads to a slow and painful death. According to the chief prosecutor, mere luck and intense police work prevented any tainted jars from being sold. Even though in court Jochen S. claimed borderline personality disorder due to his difficult childhood, a psychologist assessed the 54 years old defendant to be criminally liable despite having a narcissistic disorder. His motive was to pay off huge debts from failed business ventures. Jochen S. was sentenced to 12 years of prison on five counts of attempted murder and extortionate robbery.<sup>2</sup>
- (3) Gore sites, also known as "shock sites", enjoy a wide audience – 10 to 15 million estimated monthly visits for bestgore.com (one of the larger sites) alone – and specialize in collecting videos depicting graphic scenes of true-life killings, torture, bombings and abuse; the websites are legal in most countries and host thousands of videos. The motives for scouring the internet in search of depictions of extreme violence and suffering vary. While some of the visitors to those sites are driven by sheer curiosity and merely want to see "what's really going on", there are also those who

1 <https://www.theguardian.com/world/2018/oct/30/german-nurse-serial-killer-niels-hoegel-on-trial-100-patients-deaths>.

2 <https://www.npr.org/2018/10/23/659815368/man-sentenced-in-germany-for-poisoned-baby-food-scheme?t=1541756277147>.

watch and engage with the videos (for example by “liking” them) out of voyeuristic sadism.<sup>3</sup>

The actions of Niels Högel, Jochen S. and many anonymous visitors to gore sites can all appropriately be called evil. At the same time, each case differs structurally from the others: what stands out about Niels Högel is not just the grave harm that his actions have brought about, but the fact that he did so by using vulnerable patients for the sake of selfish goals. Like Niels Högel, Jochen S. has also used helpless human beings in pursuit of his advantage. Unlike in the first example, however, out of luck no one was actually harmed. Further, while Niels Högel must want the medication to have its hazardous effect for him to get what he wants – a dramatic life-or-death-moment that gives him a chance to demonstrate his skills – we can imagine Jochen S. being indifferent as to whether any toddler gets poisoned as long as his threat is successful. The third case also displays acts that can be called evil (even if they are less evil than the first two examples), namely watching videos for the voyeuristic pleasure of seeing people suffer and being tortured. But unlike in the first two examples, extreme harm is neither being caused nor intended by watching the videos;<sup>4</sup> furthermore, this fact is not due to luck since sadistic voyeurs visiting gore sites have no intention of causing harm by themselves, but ‘merely’ enjoy watching other people being hurt, tortured and killed.

One possible explanation of why those actions are evil has already been mentioned. According to this view, what all evil acts have in common can best be explained by reference to the reactions that we experience towards them, often termed “moral horror”. Even though such a reaction has been mentioned by a number of authors (e.g. Arendt 2007, 45; Adams 1999, 104–107; Garrard & McNaughton, 21–23; Wilson & Wilson 2003, 323; de Wijze 2019, 209–211; Kramer 2011, 220), few of them have attempted to describe its phenomenology in closer detail. Without pretension to completeness, we can gain a better grasp of what the phenomenology of moral horror consists in by considering what reactions the three examples introduced above would typically elicit. Three aspects of this reaction seem to us noteworthy.

3 <https://www.theverge.com/2012/6/13/3076557/snuff-murder-torture-internet-people-who-watch-it>.

4 This is true, at least, if the perpetrators are not attracted by expectation of internet fame (as Magnotta arguably was), causing harm and recording it for the sake of notoriety, in which case a high number of clicks might encourage the perpetrator to commit more or even graver offences. We take it, though, that the evil residing in enjoying those videos persists even if the visit to the gore site didn’t leave any trace.

First, moral horror consists at its core in a profound shock that evil actions are not only possible, but real. Even though we might know in the abstract that horrible things are happening in the world, when confronted with real evil, particularly in our surroundings, we are shaken by the sudden realization that evil is not just some distant possibility or perverted fantasy. Second, our reaction of moral horror is directed towards the evil-doer rather than towards the action or the victim; this is what makes it moral horror as opposed to more event-centered forms of dismay. At the same time, moral horror is not immediately a form of moral blame. Our reaction to an evil action is compatible with the recognition that we should, due to exempting circumstances (poor upbringing, lack of empathy etc.), suspend or at least reduce towards the agent our ordinary interpersonal attitudes and take what Strawson calls an “objective stance” (Strawson 1962). Instead, we realize that describing the action as merely wrong would not do justice to what has happened, nor that it exhausts the significance this act has for the evaluation of the agent (see also de Wijze 2019, 210). Third (a point that connects the first and the second aspect), moral horror contains an element of uncomprehending bewilderment. Hannah Arendt expresses the same point when she writes that “real evil is what elicits in us speechless horror, where we can say nothing but: this must never have happened” (Arendt 2007, 45). Evil actions therefore seem to defy a kind of understanding that is possible with regard to other actions done by humans, even highly reprehensible ones.

It is both true that a reaction with these characteristics can be elicited by something that is not in fact evil, and that what in fact *is* evil can fail to evoke it; but it is plausible to think that all and only evil actions *deserve* the reaction of moral horror. This immediately suggests a view that identifies evil with the appropriateness of moral horror; on such a view, for something to be evil *just is* for it to merit moral horror. But a fitting-attitude account of this kind will not do. For even if it is extensionally adequate – that is, even if an action is evil if, and only if, it deserves moral horror –, it gets the direction of explanation wrong. What we want to say is that an action deserves moral horror *because* it is evil, not the other way around. Conflating evil with the appropriateness of moral horror deprives us of the conceptual resources needed to account for this explanatory asymmetry.

If there are, thus, strong reasons that count against fitting-attitude accounts of evil, this raises the question of how a single theory can otherwise account for the diversity of phenomena that fall under the heading of evil actions. In response to that challenge, a number of divergent theories have been proposed.

## 2 Thick and Thin Accounts of Evil

For some philosophers, evil actions are extreme culpable wrongs. In order for an action to qualify as evil, it is irrelevant from what motive the agent has acted. While some perpetrators act out of malice, self-interest or lust for power, others such as Adolf Eichmann might – at least in Hannah Arendt’s interpretation – have acted out of banal motives and might not even be able to recognize the wrongness of their actions. Even though historical research has uncovered ample evidence to the effect that Arendt’s interpretation of Eichmann is deeply flawed (Rosenbaum 2009), her insight that evil can be committed by individuals who act without malicious intent has influenced a number of philosophers to defend “psychologically thin accounts” of evil action, as Luke Russell has called the position (Russell 2014, 79). In fact, examples that fit Arendt’s description can easily be found in many other contexts where organized forms of extreme violence occur. To mention one example, during the Indonesian mass killings of 1956–66, almost one million civilians were murdered. Many of the murderers were blind to the horribleness of their deeds and could only produce shallow forms of rationalization when later asked to justify their doings.<sup>5</sup>

Claudia Card, another defender of psychologically thin accounts, describes the view as follows. For her, evil actions are those which produce “foreseeable intolerable harms produced by culpable wrongdoing. [...] [T]he nature and severity of the harms, rather than the perpetrators’ psychological states, distinguish evils from ordinary wrongs” (Card 2002, 3).

It is a significant advantage over the more demanding psychologically thick accounts that thin accounts are able to accommodate a wide range of cases we would intuitively characterize as “evil” which stem from divergent motives or states of mind that seem difficult to unify under a single description. This advantage comes, however, at a cost: if the state of mind of the perpetrator is excluded from the analysis, we lose the ability to describe certain acts as evil that should intuitively fall under the extension of the concept. This is especially clear in the third case mentioned above: many would say that the sadistic voyeur does something evil when he gets pleasure from watching videos of other people being tortured and killed. Problems also arise from the second example. Despite the fact that Jochen S. caused no significant harm, we would not hesitate to describe his actions as evil. But other than the sadistic voyeur, he was engaged in behavior that could easily have led to the death of innocent

---

<sup>5</sup> As portrayed in the remarkable documentary “The Act of Killing” (2012).

children. Hence, his behavior stands in a closer connection to severe harm, and a defender of psychologically thin accounts might claim that this relation qualifies his behavior as evil. However, this would make the ascription of an action as evil depend on the fact that Jochen S. has *willingly* and *recklessly* exposed third parties to severe risks and hence include his psychological states. It is therefore doubtful whether thin accounts can account for such cases.

As both examples show, psychologically thin accounts are too narrow. When it comes to the above-mentioned examples, they also seem to miss something essential. There, what is evil is not just the connection to grave harm, but especially the motivation of the agents: the fact that Niels Högel poisoned helpless patients out of boredom or the fact that Jochen S. targeted small children in order to increase his chances of extracting money from the supermarkets is what makes their actions stand out. It therefore strikes us as false to claim that an account of evil can disregard the motivational setup of the agent as long as intolerable harms have been culpably produced. Instead, those psychological states of the agent should take center stage in an explanation of what makes an action evil.

This leads to so-called psychologically thick accounts, according to which a distinctive psychological state is necessary for an action to count as evil. There is, however, disagreement on how to characterize this element. Some suggest that an evil action presupposes malice (Kekes 2005), while others argue that sadistic pleasure is the psychological hallmark of evil action (McGinn 1997); other candidates include defiance of morality (Rawls 1971), the existence of a “fixed character” out of which evil actions are done (Morrow 2003) or the view according to which evil actions are those that have been chosen by agents who are blind to considerations that should silence any countervailing considerations (Garrard 1998, Garrard 2003).<sup>6</sup> The lack of convergence towards one plausible theory and the wide range of proposals that have been put forward point towards the central challenge that besets thick psychological accounts: taken in isolation, each of the proposals is open to counter-examples which demonstrate that no single psychological theory is able to cover the wide variety of motives from which evil actions can be committed.

To illustrate this problem, take the theories of McGinn and Garrard. McGinn has argued that non-instrumental sadistic pleasure is a necessary condition for an action to be evil. Non-instrumental sadistic pleasure can occur in two forms: in the first instance, somebody – e.g. a torturer or a bully – might experience pleasure by inflicting pain on her victim. Second, sadistic pleasure

---

<sup>6</sup> For a careful discussion and criticism of various forms of psychologically thick theories, see Russell 2014, ch. 4.

can be experienced by a voyeur who does not take part in causing pain, but who derives pleasure from observing it. Evil characters taking sadistic pleasure in their actions are frequently portrayed in film and literature (e.g. Joker in the later Batman movies).

While it might seem plausible that the cases that fall under his description are evil, the question is whether McGinn's account covers all possible cases where the label seems appropriate.<sup>7</sup> First, his account rules out the kind of cases Arendt's Eichmann represents, the desk-murderer who condemns his victims to death while being indifferent or morally blind. But more importantly, there are cases where the pain of others is sought instrumentally, for example when a torturer merely "does his job", but does not derive any pleasure from doing so (Garrard 2002, 328–329). It would be revisionary to exclude per definition the possibility that those actions count as evil. Indeed, McGinn's account would imply that most actions that common sense would describe as evil are not evil. From the three examples introduced above (and they could easily be multiplied), only the third would count as McGinn-evil.

An alternative version of psychologically thick accounts locates the additional component in the way agents perceive or fail to perceive certain reasons and in the way they deliberate on those reasons. Eve Garrard argues that for an action to be evil, the agent has to silence certain considerations that speak against the action. This means that the agent displays "a total failure to see that certain considerations are reasons at all, even when they are" (Garrard calls this "psychological silencing" (2002, 330)). In order for an action to count as evil, the consideration that is being psychologically silenced needs to be of a certain kind: it is such as to deprive all opposing reasons that may otherwise speak in favor of the action of their normative power (Garrard calls this "metaphysical silencing"). The reason for limiting the considerations in question to those that are metaphysical silencers is that otherwise, agents who psychologically silence trivial reasons would count as evil, for example when I disregard the reason for buying a ticket and use public transport without paying for it. Hence, evil actions are defined such that an agent psychologically silences a metaphysical silencer.

It is an advantage of Garrard's account over McGinn's that she can explain why actions as those done by Eichmann are evil: as Arendt describes it, Eichmann was aware of all relevant considerations, but engaged in psychological

---

7 There might even be cases that fall under McGinn's description where there is room for disagreement as to whether the action is evil. For example, if I make a disrespectful comment to the bus driver and draw some small pleasure from seeing him getting upset, my action is certainly wrong, but is it not too trivial to count as evil?

silencing of metaphysical silencers. But as with McGinn's theory, we can ask whether the theory covers all central cases. First, we should note how demanding Garrard's theory is. To be in a position to judge whether an action was evil, we would have to assess complex questions about the way the agent has reasoned prior to acting. In real life, it would therefore almost never be possible to categorize an action as evil; only in fiction could all the relevant information be available (Russell 2014, 100–101). Second, it is doubtful whether psychological silencing should be the watershed that divides evil actions from actions that are wrong. Imagine that in our first example, Niels Högel did not silence the reasons against poisoning his patients, but that he gave them some but not much thought. Would this make his action any less evil?

The discussion seems to have reached an impasse: on the one hand, reference to motives seems unavoidable, but on the other, none of the accounts matches our common-sense intuitions about the scope of the concept. In response, Luke Russell has argued for what he calls a "restricted conceptual pluralism" which encompasses several definitions of evil action (Russell 2014, 129; Russell 2019, 254). But before giving up on the project of finding a unified explanation of evil action, it might be worth exploring the possibility of a psychological thick account which is able to allow on its own terms for various ways in which the mental condition can be spelled out. This is what our account aims to achieve.

### 3 The Manifestation Account of Evil

In a nutshell, the definition of evil action we propose is the following: an action is evil if and only if, and because, it is (1) the manifestation in action of (2) a pro-attitude (or complete indifference) towards severe harm to a sentient being, and (3) this attitude is either spontaneous or the outcome of a broader bad attitude. Of these conditions, (1) refers to the way(s) in which a mental state of an agent manifests itself in outward behavior, (2) to that mental state itself, and (3) to the conditions that give rise to the mental state; accordingly, we will refer to these conditions as "behavioral", "mental", and "causal". By acknowledging a mental condition, we subscribe to a thick, psychologically rich conception of evil. On our account, evil actions differ from ordinary wrongdoings not only in respect of the severity of harm done, but also regarding the mental state of the perpetrator. However, since neither the mental state nor the condition from which it issues have to be stable traits of character, we also reject the view that the notion of evil personhood is prior to that of evil actions. In our view, a person is evil if she is prone to do evil actions, so that evil action can

and must be understood independently of what it is to be an evil person. In what follows, we will discuss each of the three elements of our account.

### 3.1 *Evil Actions as Manifestations*

Assuming that a pro-attitude towards harm is necessary, in what relation to that pro-attitude does an action have to stand in order for it to be evil? The mere presence of the mental element is obviously not enough. I might be possessed by an intense desire to harm another person, but if there is a stronger attitude holding me back, my pro-attitude towards the harm may never show itself in the way I act and my acting will not be evil. One might hold that in order to count as evil, an action must be done *in order to satisfy* the underlying pro-attitude. This is a tempting but misleading thought. While some actions are indeed evil in virtue of bringing about a certain effect (harm) that the agent has a pro-attitude towards, others are evil in virtue of expressing some pro-attitude such as joy at a certain harm.<sup>8</sup> Thus, the wanton killing of an innocent family described in Truman Capote's *In Cold Blood* is evil because it realises the protagonists' blood lust; Eichmann's statement that he was happy to have brought "five million Jews to the grave" was evil because it expresses his satisfaction at the mass killing of innocent human beings.

A unifying rationale behind these two ways is provided by the idea that pro-attitudes stand to actions in the relation of dispositions to their manifestations. Following Hyman, a pro-attitude (or "desire" in the technical, philosophical sense) can be manifested in essentially two ways (Hyman 2015, 107–109). First, a pro-attitude can be manifested in goal-directed behavior, i.e. behavior aimed at bringing about a certain state of affairs; second, a desire can be manifested in the way a certain emotional reaction to an object gets expressed. For instance, the child's desire for milk can be manifested by its trying to grab the cup; but the same desire can also be present in the joy on the child's face as it is given the cup of milk, or in its protest as the milk is taken away from it.

For behavior to express a certain desire, the agent doesn't have to be aiming at that expression (he just expresses it); however, it is of course perfectly possible that he does aim at expressing his desire. In the same vein, an action can manifest a pro-attitude toward the harm for other beings in either of two ways: either by goal-directed behavior that aims at bringing about the harm or by expressing an emotional reaction towards its occurrence (or non-occurrence).

The distinction between two kinds of manifestation allows us to take on board a number of cases that would otherwise be difficult to accommodate. For instance, we can now say that what's wrong with watching gore videos, making

---

<sup>8</sup> For a discussion of expressive action, see Hursthouse 1991.

nasty comments about the death of other people etc. is not only that this leads to further suffering or that it is the viewer's intention to bring this about. What makes such actions evil is that they *express* their agent's pleasure at, and improvement of, the suffering that is done to other people. Likewise, in Garrard's case of a tyrannical state charging the grieving relatives of a young dissident for the cost of the bullet that was used to execute him (Garrard 1998, 45), what strikes us as evil is not so much the harm that is wantonly inflicted upon the relatives but the cynical expression of indifference towards the young man's suffering. In a similar vein, Kramer speaks of the "utter disdain for the humanity and elementary well-being of someone else" expressed in the action (2011, 216).

### 3.2 *The Pro-attitudes behind Evil Actions*

What exactly is the pro-attitude that is manifested in an evil action? What do we mean by "pro-attitude" in general, and what are evil pro-attitudes directed towards? Finally, is the object in an evil pro-attitude necessarily desired for its own sake, or can it also be desired for the sake of something else?

#### 3.2.1 The Concept of a "Pro-attitude"

Roughly, we have a pro-attitude whenever we want something. As Donald Davidson has defined it, pro-attitudes (or desires in the wide, philosophical sense) encompass not only desires (in the narrow, non-philosophical sense), but also "wantings, urges, promptings, and a great variety of moral views, aesthetic principles, economic prejudices, social conventions, and public and private goals and values, in so far as these can be interpreted as attitudes of an agent directed toward actions of a certain kind" (Davidson 1963, 686). That is, pro-attitudes are not restricted to states with the characteristic feel of an inner urge (such as the "yen to drink a can of paint" or simple thirst) but include action-guiding normative judgements, arguably lacking any such felt quality. As a related point, these mental states can vary widely concerning their stability and temporal extension, from "permanent character traits that show themselves in the lifetime of behavior" to "the most passing fancy that prompts a unique action" (Davidson 1963, 686).

This comprehensive notion of a pro-attitude does justice to the fact that evil actions can be done from a wide range of different mental states, from the rage of the moment to careful and extended deliberation. In particular, it means that other than defenders of the "fixed conception of evil" argue (Russell 2006, 97–99; Morrow 2003, 234; cf. Garrard 1998, 44), evil actions can, but don't have to, be done out of a fixed character; instead, they might also be the result of yielding to a sudden desire. This means that for one's action to be evil, one

doesn't have to be an evil person, if by this we understand a person with a stable disposition to act in evil ways. To illustrate this, consider the case of Svetoslav Stoykov, who in an underground station in Berlin kicked a random woman down a flight of stairs and then nonchalantly walked away with his companions.<sup>9</sup> Even without knowing whether his action was the result of long premeditation or whether Stoykov acted out of a sudden impulse, it is clear that what he has done counts as evil.

### 3.2.2 The Object of the Pro-attitudes

The pro-attitudes underlying evil actions are directed towards severe harm to a sentient being. Strictly speaking, the pro-attitude can be directed either towards severe harm or towards the violation of human dignity. While violations of dignity tend to also impair the victim's well-being, this is by no means necessarily the case. It is only for the sake of brevity that we omit mention of dignity violations.

While the suggestion that the underlying pro-attitudes of evil action must be directed towards severe harm has intuitive plausibility, one might wonder why it excludes forms of harm that are not severe. Can reprehensible pro-attitudes not just as well be directed towards less-than-severe harm? Further, it needs to be explained why our account excludes non-sentient beings. Shouldn't we allow for the possibility that evil actions can also be directed against, e.g., artefacts?

It would certainly not be linguistically inappropriate to call some actions that damage artefacts evil. However, the fact that the English language distinguishes between "harm" or "injury", which applies to sentient beings, and "damage", which standardly is applicable to non-living things, already indicates that there is a categorical difference to be taken into account. Correspondingly, the feeling of moral horror is only appropriate when we witness or learn about an agent who acts out of a pro-attitude towards severe harm to sentient beings. As a possible counter-example to this claim, the destruction of cultural artefacts, such as the destruction of Palmyra by ISIS, strikes many as not only barbaric, but downright evil. But unlike trivial acts of vandalism such as the destruction of a bus-stop, the destruction of Palmyra expresses hatred against other cultures and their members and should hence be classified as evil on account of the underlying pro-attitude towards their severe harm.

What, then, makes harm severe? Our suggestion is that harm is severe if it precludes someone from being in an overall state of flourishing. If we twist our

<sup>9</sup> See <https://www.theguardian.com/world/2016/dec/19/suspect-arrested-after-berlin-underground-attack-on-woman>.

colleague's arm, that will be temporarily painful, but it does not preclude them from being in a good state, all things considered. If someone is raped, however, this will not just be a partial subtraction to their well-being; it is apt to overshadow their entire life. This is why genuinely evil acts have a destructive depth that mere wrongdoings, or minor harms, lack.

At the other end of the spectrum, one might wonder why cases of less than extreme harm should not also be candidates for being evil. There are doubtlessly mean pranks such as placing a pushpin on somebody's chair for one's own amusement. While this hurts, it would be a stretch to call such actions evil; they also fail to evoke moral horror. Borrowing a distinction Kant draws between actions that are *böse* (evil) and such that are merely *boshafft* (malicious), actions that manifest a pro-attitude towards less-than-severe harm should be classified as malicious, not evil. This allows us to exclude so-called "small-scale evils" from the extension of evil proper, while acknowledging the continuities between both.

### 3.2.3 Final and Derivative Pro-attitudes

It may be thought that an action is only truly evil if severe harm to a sentient being is affirmed for its own sake (call this a "final pro-attitude towards harm"). Restricting the concept of evil in this way meets a pervasive cliché of Western culture, from Milton's Satan ("Evil, be thou my good") to his latter-day successors such as Moriarty or the Joker. However, stipulating that the perpetrator must want the harm for its own sake if the corresponding action is to count as evil would have drawbacks. It is disputed whether acting for the sake of evil is conceptually possible, leaving the extension of the concept potentially empty.<sup>10</sup> But even if it is possible to act under the guise of the bad – a question on which we remain agnostic, but which does not raise problems for our definition either way – it would be a rare occurrence, as the fact that all three of the examples introduced above would not fall into this category indicates: Niels Högel acted out of boredom and craving for recognition, Jochen S. in order to extort money from retailers, and the nameless visitor to the gore site watches the videos in order to take delight in other people's suffering. Not all evil is satanic evil.

We should therefore widen the extension of the pro-attitudes potentially underlying evil actions to include derivative pro-attitudes. The simplest case is when we have a pro-attitude towards something because it is the necessary

<sup>10</sup> While for Ockham, doing evil for its own sake is possible (cf. *Connex* 3,52148, cf. *Sent.* 1.38; for discussion, see McCord Adams 1999, esp. 255–261), Aristotle argues that one cannot choose an action for an evil end (cf. *NE*, 1,1 and Korsgaard 2008, 149).

means to, or the unavoidable side-effect of, something else we want. This is so in the cases of Högel and Jochen S., who both welcomed or accepted the harm done to their victims for its instrumental relation to their more fundamental ends. Likewise, the mass killings of the Jews in Nazi Germany were justified by the alleged higher aim of Germany's "racial purity". But there are other kinds of derivative pro-attitudes beyond the instrumental. In particular, it is possible to take what can be called "constitutive pro-attitudes", where the harm is desired for the sake of something else, while this, in turn, cannot be described without recurring to the harm itself. Sadistic pleasure falls in this category. To the sadist, the harm and suffering of her victim is not just a means to obtain generic pleasure that she might feel equally well by doing other pleasurable things. The harm is built right into the pleasure; it is pleasure-in-harm, so to speak. So while the infliction of harm is necessary for the achievement of the pleasure, the pleasure cannot be specified independently of the harm; it is, in part, constituted by it. The fact that the pro-attitude behind evil actions is by no means necessarily ungrounded, but can be present for a variety of reasons, also explains why from the perspective of the agent herself the evil action can appear completely justified.

To sum up, we intend a broad understanding of pro-attitudes that can underlie evil actions. Not only can severe harm to sentient beings be wanted either for its own sake (final) or for the sake of something else to which it stands in a suitable relation (derivative), the pro-attitude can also be either transient or lasting, either "hot" (emotionally charged) or "cool" (resulting from calm deliberation). "Harm" is likewise meant to be comprehensive, encompassing both the dimension of well-being and that of fundamental rights. Further, for an action to be evil, the harm to a sentient being that the pro-attitude is directed towards must be severe.

### 3.3 *The Attitudes behind the Pro-attitudes*

Although a pro-attitude with these characteristics must be present, this is not sufficient for an action to count as evil. To see why, take the case of an executioner, Joe, working within a largely legitimate liberal-democratic system such as the United States. If he is convinced that the convict deserves the death penalty, and that he has been sentenced to death in accordance with the relevant legal procedures, he may be said to have a pro-attitude towards severe harm to a sentient being. When he actually carries out the execution, this pro-attitude will manifest in goal-directed action. Still, Joe can't be said to do something evil. He may – depending on one's stance towards the death penalty – be doing something profoundly wrong, but his action should not count as *evil*. A natural approach would be to augment the behavioral and mental conditions

by adding a clause that in order for an action to be evil, the desired harm must be undeserved.

This will not do either, as can be demonstrated by means of a second familiar example, borrowed from Bernard Williams (1973, 97): Jim finds himself in the central square of a small South American town. Tied up against the wall are a row of twenty Indians, in front of them several armed men in uniform. The captain offers Jim the “guest’s privilege” of killing one of the Indians himself; if he accepts the invitation, the captain promises to set the others free. Otherwise, all twenty will be killed. Even if we assume that shooting the one Indian would be the wrong thing for Jim to do, it would not be evil. As above, Jim’s action manifests a pro-attitude towards a severe harm to a fellow being. But unlike in the executioner’s case, none of the Indians have done anything to deserve the harm. To the contrary, the harm Jim is asked to inflict onto the one Indian will be a mere means to save the other nineteen. Hence, it is not enough to add to the mental and the behavioral condition that the harm must not be deserved. But what is it that stops Joe’s and Jim’s actions from being evil?

An option would be to locate the common feature in the fact that both Jim and Joe take their actions to be (morally) justified, whether or not that belief is correct. From this we might be tempted to infer that it is a necessary condition of evil actions that the action is, even by the agent’s own lights, morally unjustified. But then a number of cases will escape us where the agent takes his action (by his own perverse standards) to be completely justified, and yet we wouldn’t hesitate to call it evil. Take as an example Himmler’s Posen speech of 1943, where he justifies the crimes committed by the members of the SS by reference to the “higher purpose” they help to bring about, or Anders Breivik’s mass killings on the island of Utoeya, defended at length in his 1.518 pages manifesto. Taking into consideration only whether the agent saw his action as morally justified won’t work either.

Another option for the additional requirement is the normative situation as it objectively is, not what the agent takes it to be – i.e., what counts is the normative situation based on what Scanlon calls the “reasons in the standard normative sense” (Scanlon 1998, ch. 1). Of course, that we have overall reason to refrain from the action is too weak, for that may be true of Joe and Jim, too, without making their actions evil. But we could think that in evil actions, very strong reasons speak against the action, either vastly outweighing or even silencing competing considerations. For instance, Niels Högel acted for reasons that have no weight at all in the light of the (silencing) reasons against poisoning his patients. Likewise, Jochen S. seems to neglect aspects of his action (the exposure of little children to the risk of a painful death) that not only speak decisively against it but make anything that could be said for it barely count at

all. The same could not be said about Jim or Joe: although in both cases overriding reasons may speak against killing the one Indian or against conducting the execution, those reasons do not silence all competing considerations. There are considerations that speak in fact in favour of killing the one Indian (namely, the fact that by doing so, nineteen would be saved); and there is a case to be made for the idea that the death penalty is the only appropriate way of responding to certain crimes. This also nicely captures why there is something incomprehensible about genuinely evil actions (“how could he fail to see this?”), while no such incomprehensibility attaches to the actions of Joe and Jim.

Imagine, however, Jim’s evil counterpart. Evil Jim is a racist who goes to South America with the firm intention to pick off some Indios, should the opportunity arise. When he is offered by the captain the “guest’s privilege”, he is delighted: he can fulfil his long-held desire and even be on the safe side morally! Upon examination, he will be capable to justify himself by reference to the other lives he has saved. Nevertheless, Evil Jim’s action should be classified as evil, while Good Jim’s action may well be wrong, but not evil. Hence, the difference between the two cannot be accounted for by reference to the normative situation as it objectively is. But what, then, is it that makes Evil Jim’s action evil?

The answer must have to do with the way the pro-attitude towards severe harm for a sentient being has been formed: after all, Evil Jim is a racist, while Good Jim honestly tries to do the right thing. So the crucial factor seems to be where that desire comes from. This does, however, not mean that the pro-attitude has to be carefully premeditated. Even if Evil Jim does not have a fixed racist disposition and does not go to South America with the intention to kill someone, even if his desire is a matter of a sudden impulse, his action would still be evil. Therefore, the origin of the pro-attitude that makes it bad (and the corresponding evil action) doesn’t have to be (although it can be) something as stable and settled as a fixed character trait.

Rather, it is a more general (even if short-lived) attitude that generates the pro-attitude to harm which determines whether the action that manifests the latter counts as evil. There is a number of morally reprehensible attitudes that can issue in such pro-attitudes. Here is a selection: xenophobia, avarice, selfishness, anti-Semitism, vanity, craving for admiration, nationalism, sex addiction, misogyny, homophobia, arrogance, and lust for power. All these, and doubtless countless others, can, given suitable circumstances, give rise to a pro-attitude towards severe harm to others, and if that pro-attitude gets manifested in action, the resulting action will be evil. Jochen S. was driven by greed, and this led to his recklessly putting the lives of little children at danger. Niels

Högel accepted the death of innocent patients out of his desire for diversion and recognition. The sadistic voyeur's pleasure in watching torture videos is a result, or a specification, of his more general pro-attitude towards atrocities. On the other hand, a desire for justice (if it is not confused with a desire for revenge) or a benevolent attitude towards other people, although they can at occasion (as the cases of Joe and Jim demonstrate) also lead to pro-attitudes towards severe harm to others, never make the corresponding actions evil. This is true even if the process that leads to the formation of the desire for harm is deeply flawed and the resulting action seriously wrong, as possibly in the case of Jim and that of Joe.

We therefore conclude that the crucial difference between evil and non-evil actions resides in the general attitudes from which the more specific pro-attitudes towards harm arise. Given the huge variety of "bad" attitudes that can in the final instance lead to evil actions, we won't offer a general theory of what those attitudes have in common. Such a task belongs to a substantive theory of evil and falls outside the scope of a definition of the concept.

It is important to note that more often than not, those reprehensible attitudes will find expression in pro-attitudes that don't make the resulting actions evil, e.g. when they are desires for certain goods for oneself (or for less-than-severe harm to others). Also, we don't take a stand on the question whether those attitudes themselves, or the pro-attitudes issuing from them, deserve to be called evil. All we have been interested in in this article is what makes actions evil. Interestingly, this can't be read off from the actions themselves.

#### 4 Conclusion

We have defended a tripartite account of evil action. The action (observable behavior) must be the manifestation of a pro-attitude towards severe harm to a sentient being, with a broad understanding of both "pro-attitude" and "harm". The manifestation can take either of two forms: expressing the pro-attitude and attempting to realise its object. However, not every such action will be evil. This is because the occurrent pro-attitude can be the result of a positive attitude such as love for justice or benevolence. Therefore, the pro-attitude behind an evil action must itself be the result of a more general negative attitude such as greed, sadistic pleasure, or drive for recognition.

This account is flexible enough to account for the wide diversity of actions that we are inclined to call evil. While acknowledging that the connection to severe harm is not accidental to evil actions, it insists that not every evil action is so in virtue of bringing about massive harm (such as Niels Högel) or in virtue

of accepting harm as a side-effect of acting (such as Jochen S.). Rather, an action can be evil in virtue of expressing the agent's pro-attitude towards severe harm as in the case of the sadistic voyeur, or in fact in the case of certain hate comments on the internet.

Employing the concept of a pro-attitude allows us to account for a wide variety of motives from sadistic pleasure to indifferent disregard; that it has to be a pro-attitude towards harm will easily explain why evil actions are necessarily wrongdoings. But it does not follow that each and every evil action is also extremely bad, as the case of the tortured housefly demonstrates: although this constitutes severe harm to a (presumably) sentient being, there are non-evil wrongdoings that are probably much worse, e. g. negligently putting at risk the life of one's child. Evil therefore doesn't necessarily occupy the "red zone" of moral badness.

### References

- Adams, Marilyn McCord 1999a. "Ockham on Will, Nature, and Morality." In: *The Cambridge Companion to Ockham*. Edited by Paul Vincent Spade, Cambridge: Cambridge University Press, 245–272.
- Adams, Robert Merrihew 1999b. *Finite and Infinite Goods. A Framework for Ethics*. Oxford: Oxford University Press.
- Arendt, Hannah 1978. *The Jew as Pariah*. New York: Grove.
- Arendt, Hannah 2006. *Eichmann in Jerusalem*. New York: Penguin.
- Arendt, Hannah 2007. *Über das Böse. Eine Vorlesung zu Fragen der Ethik*. Edited by Jerome Kuhn, translated by Ursula Ludz, München: Piper.
- Card, Claudia 2002. *The Atrocity Paradigm*. New York: Oxford University Press.
- Clendinnen, Inga 1999. *Reading the Holocaust*. Cambridge: Cambridge University Press.
- Davidson, Donald 1963. "Actions, Reasons, and Causes." *The Journal of Philosophy* 60, 685–700.
- de Wijze, Stephen 2019. "Defining the Concept of Evil: Insights from Our Pre-cognitive Responses." In: *The Routledge Handbook of the Philosophy of Evil*. Edited by Thomas Nys & Stephen de Wijze, New York: Routledge, 203–217.
- Gaita, Raimond 2000. "Evil beyond Vice." In: *A Common Humanity: Thinking about Love and Truth and Justice*, London & New York: Routledge, 29–56.
- Garrard, Eve 1998. "The Nature of Evil." *Philosophical Explorations* 1, 43–60.
- Garrard, Eve 2002. "Evil as an Explanatory Concept." *The Monist* 85, 320–336.
- Garrard, Eve & David McNaughton 2018. "How to Theorize about Evil". In: *Moral Evil in Practical Ethics*. Edited by Shlomit Harrosh & Roger Crisp, Oxford: Routledge, 7–29.

- Hanfling, Oswald 2003. *Philosophy and Ordinary Language: the Bent and Genius of our Tongue*. London: Routledge.
- Haybron, Daniel 2002. "Moral Monsters and Saints." *The Monist* 85, 260–284.
- Hursthouse, Rosalind 1991. "Arational Actions." *The Journal of Philosophy* 88, 57–68.
- Hyman, John 2015. *Action, Knowledge, and Will*. Oxford: Oxford University Press.
- Kekes, John 2005. *The Roots of Evil*. Ithaca, N. Y.: Cornell University Press.
- Korsgaard, Christine M. 2008. *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*. Oxford: Oxford University Press.
- Kramer, Matthew H. 2011. *The Ethics of Capital Punishment: A Philosophical Investigation of Evil and its Consequences*. Oxford: Oxford University Press.
- McGinn, Colin 1997. *Evil, Ethics and Fiction*. Oxford: Clarendon Press.
- Morrow, Lance 2003. *Evil: An Investigation*. New York: Basic.
- Neiman, Susan 2003. *Evil in Modern Thought*. Melbourne: Scribe.
- Rawls, John 1971. *A Theory of Justice*. Oxford: Clarendon Press.
- Rosen, Gideon 2015. "Real Definition." *Analytic Philosophy* 56, 189–209.
- Rosenbaum, Ron 2009. *The Evil of Banality*. url: [www.slate.com/articles/life/the\\_spectator/2009/10/the\\_evil\\_of\\_banality.html](http://www.slate.com/articles/life/the_spectator/2009/10/the_evil_of_banality.html).
- Russell, Luke 2006. "Evil-Revivalism Versus Evil-Skepticism." *The Journal of Value Inquiry* 40, 89–105.
- Russell, Luke 2007. "Is Evil Action Qualitatively Distinct from Ordinary Wrongdoing?" *Australasian Journal of Philosophy* 85, 659–667.
- Russell, Luke 2014. *Evil: A Philosophical Investigation*. Oxford: Oxford University Press.
- Russell, Luke 2019. "Defining Evil Actions: Different Approaches." In: *The Routledge Handbook of the Philosophy of Evil*. Edited by Thomas Nys & Stephen de Wijze, New York: Routledge, 245–255.
- Scanlon, T. M. 1998. *What We Owe to Each Other*. Cambridge, MA: Belknap Press of Harvard University Press.
- Singer, Marcus G. 2004. "The Concept of Evil." *Philosophy* 79, 185–214.
- Strawson, Peter 1962. "Freedom and Resentment". *Proceedings of the British Academy* 48, 1–25.
- Williams, Bernard 1973. *Utilitarianism: For and Against*. With J. J. C. Smart. Cambridge: Cambridge University Press.
- Wilson, John & Nicholas Wilson 2003. "Moral Horror." *Journal of Beliefs & Values* 24, 321–327.