

Kantian Rigorism Revised

A defense of Schapiro's revised Kantianism as a response to the case of the murderer at the door



Utrecht University

Department of Philosophy & Religious Studies

Abstract: In a notorious essay, Kant defends that one is obligated to tell the truth to a murderer who asks where his victim is. This has been discussed in literature as ‘the problem of rigorism,’ which supposedly does not allow for mitigating circumstances. Several responses exist to this problem, which can be distinguished as rigorism, Kantian apologetics, structural Kantianism, and revised Kantianism. In this thesis, I investigate these strategies and defend Schapiro's revised Kantianism, because it allows for non-ideal situations to mitigate the stringency of the categorical imperative whilst maintaining the core tenets of Kant's deontology. In doing so, I refute Cholbi's anything-goes objection to her constitutive approach.

Keywords: Kant · Deontology · Rigorism · Lying · Non-ideal theory

By:	Peter Saarloos
Supervisor:	dr. Jan Vorstenbosch
Secondary supervisor:	dr. Jos Philips
Word-count:	20,298 (+3,770 footnotes & references)
Date:	December 22, 2017

Index

Introduction	3
Chapter 1: On Kant's rigorism	8
1.1. The categorical imperative and perfect duties	8
1.2. Rigorism	10
1.3. Kant on the supposed right to lie	13
1.4. Kantian apologists	16
1.5. Paton on exceptions to perfect duties	18
Chapter conclusion	21
Chapter 2: Structural Kantianism	23
2.1. Cholbi on the duty of the preservation of others	23
2.1.a. Cholbi on perfect duties	25
2.1.b. Deontic symmetry	26
2.1.c. Refuting objections	28
2.2. Response to Cholbi	30
2.2.a. Cholbi and non-ideality	31
2.2.b. Cholbi's deontic symmetry thesis	33
Conclusion of this chapter	35
Chapter 3: Schapiro's revised Kantianism	36
3.1. Problems of the standard argument	36
3.2. Schapiro on honesty: letter and spirit	38
3.3. Cholbi's 'anything goes' objection	41
Chapter conclusion	42
Chapter 4: Revised Kantianism	44
4.1. Social practices and non-ideal conditions	46
4.2. The spirit of the moral law	50
4.3. Concerns and further discussion	54
Conclusion	56
References	59

Introduction

Kant's deontology is a moral theory with strong intuitive appeal that is very functional in ordinary moral considerations. Its morality respects the intuition that there are certain things we just shouldn't do, contrary to the theory's main competitor consequentialism, which in principle forbids no action for reasons inherent to the nature of the act. Additionally, deontology provides a clear theoretical framework for reasoning about morality, contrary to relativism or moral skepticism, but also contrary to some other moral theories other than consequentialism, which struggle to draw clear conclusions in everyday moral practice. Once we accept the premises of Kant's morality, the autonomy of moral agents, and we are persuaded by the validity of the categorical imperative, we enter into a realm of moral deliberation in which there is a clear framework for permissible and impermissible action.

Yet Kant's moral theory is not uncontroversial. Much of this controversy is derived from the ties it has to Kant's entire body of work. Kant was eager to create a complete and coherent philosophy and brilliantly succeeded in doing so: his epistemology allows space for his ethics, generating the metaethical possibility of a morality quite naturally (i.e. the postulate of the soul). Yet the practical limits of that morality seem to trouble Kant, and his eagerness to create a systematic morality generated problems that he could not adequately respond to in his lifetime.

To create a morality clearly delineated from his natural philosophy, Kant separates free action, which humanity is capable of insofar as they are rational and autonomous agents, from mere behavior caused by inclinations, which influence humans insofar as they are natural beings. This leads to a position called rigorism, which Kant forwards in *Religion within the Boundaries of Mere Reason*. There, Kant argues that an action is either good or evil, and that an agent is therefore always good or evil as well, depending on whether or not she acted out of natural inclination or in accord with the moral law. The moral law consists of the categorical imperative and those maxims that withstand to be tested against it and are therefore moral laws. This means that our practical circumstances are rendered secondary to that test: if a maxim fails to be validated by the three formulations of the categorical imperative, it is impermissible, and the intuition that it should really be permissible is reduced to mere inclination. The latter means that bad actions are the result of self-love that

corrupts our ability to reason about what is right.¹ Good actions, therefore, are those actions that are obligatory or at least permissible under the moral law, and any agent that does not test his maxims against the moral law adequately is doing so out of vice, and is therefore evil. There is, according to Kant's rigorism, no middle ground or a lesser of two evils

In academic literature, this rigorism is considered problematic, and a discussion has emerged about rigorism and the possibility of mitigating circumstances. The case that is discussed in light of this problem is that of *the murderer at the door*: if a man is at your doorstep and asks you if you are hiding his victim (him naïvely assuming that you will cooperate), are you allowed to deceive him? Prima facie, the maxim of deceiving another agent fails twofold in light of Kant's formulations of the categorical imperative: first, deception cannot be universalized, because we cannot imagine a world in which anyone lies and deceives when it suits them; second, deception runs directly counter to the second formulation of the categorical imperative, which states that we should treat others as ends and not as means (known as the *kingdom of ends* formulation). Kant's own words on the matter in his essay *On the Supposed Right to Lie from Philanthropic Motives* are that an agent simply shouldn't lie in that circumstance, which conforms with his rigorism.

This case, at least in the view of myself and the scholars involved in the discussion, is not trivial. If my previous praise for Kantianism is warranted in that it is an intuitive yet sufficiently action guiding moral theory, that is a great asset against moral skeptics and opens the possibility of systematic work in moral theory. Yet it would be detrimental to the theory if it should fail precisely when it ought to be informative, namely in dilemmatic situations. Thus the case becomes a symptom for a more general problem, namely of the relationship between the categorical imperative and its formulations and the way in which agents are informed by this. It cannot be a satisfying answer that Kantianism works as long as it overlaps with our intuitions, but that we should abandon it whenever its conclusions are counter-intuitive: that is precisely *not* the point of a normative theory, and would render its mode of reasoning mere rationalization.² Thus the case has really served as a

¹ This kind of self-love, which is natural inclination, ties back into his epistemology, because it is this part of human action (or, more properly, human behavior) that is causally determined – it is morality that is properly free human action, because it frees itself precisely from that natural inclination.

² On the contrary, it seems the case that when a theory is entirely counter-intuitive, it wouldn't even be worthy of consideration as a moral theory, and it seems that there is a way in which our

discussion regarding what happens when the practical implications of Kant's moral theory are pressed.

Aim of the thesis

In a recent article, Cholbi criticizes Schapiro's solution for the former case. She argues for a constitutive approach which emphasizes the relational aspects of honesty and the moral law, what she calls the moral community. The murderer, in intending to murder, betrays this moral community and thus the agent is allowed to act from a position of refusal, excusing the lie. Cholbi rejects this approach, because it is unclear that it is precisely a lie that is permitted in this conception. If the murderer betrays the moral community, it seems that anything goes as a result.

Yet in an earlier article, Schapiro discusses non-ideal circumstances and social practices, specifically in the case in which social practices are usurped by some agent who uses the practice to further his personal ends. In this article, she distinguishes two ways in which an agent can relate to a social practice, namely as social engineer and as reformer. The social engineer thinks of practices as a mere method to achieve some end, whereas the former wishes to be both a participant in the practice as well as a restorer of the practice in face of its corruption (due to it being usurped). At the end of the article, she expresses the ambition to investigate the possibility of extending this conception to morality at large, questioning whether the moral community and honesty can be conceived of as practices in a similar vein.

Cholbi's criticism can be responded to once we consider what the role of the reformer might look like in the context of the moral community at large. The aim of this thesis is therefore to apply the insights of Schapiro's former article to the moral community at large as a response to Cholbi and, in the process, gain insight into how the constitutive approach to Kantianism deals with non-ideal circumstances. Once we are able to understand the case of the murderer of the door as a case of usurping the moral community and once we understand the role of the agent in this framework, the anything-goes objection can be deflected.

What is pivotal in settling the matter of rigorism is that sense can be made of non-ideal situations in a manner that makes intelligible the dilemmatic character of such

moral reasoning is intimately related to our intuitions on the matter. This is precisely what makes consequentialism seem absurd at times, such as when it suggests that practices like organ harvesting are perfectly acceptable means if it generates sufficient moral good.

circumstances. Agents must on the one hand aim to act in the spirit of morality, but on the other hand recognize that non-ideal circumstances disallow them from complying fully. If this troubled position of the agent can be made sense of without succumbing to the anything-goes objection, then Schapiro's revised Kantianism is on steady ground in providing a satisfying response to rigorism and the case of the murderer at the door.

Structure

An important question in the topic of rigorism is whether Kantian moral theory as classically conceived is equipped to deal with the case of the murderer at the door. To answer this question, four types of responses to the matter will be considered. The first response is simply embracing the conclusions drawn by Kant in *On the Supposed Right to Lie*, and proponents of this view may therefore be called **rigorists**. I will elaborate on this position in chapter 1, and although my purpose is not to offer a historic interpretation of Kant, I aim to at least make rigorism plausible by discussing his work. A second response consists of **Kantian apologists** who consider the problem to arise primarily because the essay exists in the first place, and that once the essay is rejected Kant's moral theory is equipped to deal with the matter in itself.³ I disagree with this view, not necessarily because I think my interpretation of Kant is the correct one or that the essay shouldn't be rejected, yet my aim with chapter 1 is to show that the case of the murderer at the door is a real problem derived from the formulations of the categorical imperative. I will therefore discuss arguments for and against Kantian apologists in the same chapter.

The third type of answer is what I will call **systematic Kantianism**. It is a view that accepts that rigorism may form a problem for Kant's moral theory, and while rejecting the conclusion drawn in *On the Supposed Right to Lie*, it seeks to develop a systematic solution to rigorism based on Kant's work as well as the work in the Kantian tradition. Systematic Kantianism does not believe that Kant necessarily provided the answer to the problem of rigorism and systematic work may be required to complete his view, but systematic Kantianism does believe that the answer can be derived from the formulations of the categorical imperative provided by Kant. In the recent discussion on the topic, Cholbi is a defendant of systematic Kantianism, and I will discuss his arguments in chapter 2.

³ I take the term 'Kantianism' to refer to any philosophy that is inspired by Kant or is part of that deontological tradition, but that is not the work of Kant himself.

The fourth type of response seeks to offer a revised view of the formulations of the categorical imperative because it considers the problem of rigorism to be derived from Kant's own conception of the categorical imperative. I will therefore call this view **revised Kantianism**, and in the recent discussion Schapiro is a proponent of it.

Throughout this discussion the case of the murderer at the door serves as a placeholder for the more wide consideration of what Schapiro calls 'defensive deception' and, more broadly, on the relationship between intuition and rigorism in general. It cannot be the case that we answer the particular case as sketched in the discussion, but fail to develop a methodology that allows us to answer such cases more generally, unless the murderer at the door were the only case in which Kantianism is troubled by these tensions. It is not: Russell mentions the case of breaching a contract to construct a dam in order to save the water supply of a village and potentially countless lives. A breach of contract, like deception, is difficult to justify in terms of Kant's formulations of the categorical imperative, and although perhaps a satisfying rationale can be developed for this particular case, practical cases should not become piles of paperwork for Kantians when our moral intuitions on the matter are completely clear. Therefore the way in which this case is handled is telling for the entire methodology of Kantianism, and thus the answer can only be satisfying if it has somehow deepened our understanding of Kantian deontology and its operation in practice.

Chapter 1: On Kant's rigorism

The discussion about the significance of the case of the murderer at the door and the problem of rigorism has a tradition in literature. One approach in that discussion, particularly forwarded by Paton and Benton in their respective articles “An Alleged Right to Lie” and “Political Expediency and Lying: Kant vs Benjamin Constant”, was to forward an interpretation of Kant in which the essay is ignored or qualified for historic reasons.⁴ The idea is that, when the essay is removed from the discussion, the issue of the case and the greater problem of rigorism can be resolved with ease simply by referring to Kant's original work. In the present chapter, it is my primary aim to demonstrate that this is not necessarily the case. What gives the discussion substance is that rigorism and its problematic aspects do follow quite naturally from Kant's original moral philosophy. This is not to say that Kant meant to be as rigoristic as he is in *On the Supposed Right to Lie* – Paton and Benton may be right in suggesting Kant had other motives for concluding what he did in that essay. My aim however is to show that there is a real problem that follows from the formulations of the categorical imperative offered by Kant, and that we need to revise them in order to make them sufficiently action guiding.

In the first part of this chapter, I will therefore discuss several elements of Kant's moral philosophy that I think contribute to this view. In the second part, I will respond to Paton's philosophical arguments for ignoring the conclusions of the essay (and not his historical ones).

1.1. The categorical imperative and perfect duties

It was Kant's project explicitly to investigate the possibility of an a priori moral theory in his *Groundwork for the Metaphysic of Morals*.⁵ The aim of the project is to marry rationality and morality, i.e. to be able to conclude reasonably which actions are permissible, obligatory, or forbidden. Actions based on natural impulses are therefore

⁴ Paton, “An Alleged Right to Lie,” *Kant-Studien* 45, 1-4 (1954): 190-203 and Benton, “Political Expediency and Lying: Kant vs Benjamin Constant,” *Journal of the History of Ideas*, Vol 43, No. 1 (Jan. – Mar., 1982): 135-144.

⁵ See his description of the project at G 4:389, 5. I take this to be the reason of Kant's popularity and his relevance in moral philosophy: he conceived of the possibility of a moral theory as the rules to the game of life and how to play it fairly. Regardless of its problems, this idea is brilliantly revolutionary and it should be no surprise that its possibility has inspired many philosophers since.

Citations from Kant, *Groundwork of the Metaphysics of Morals: A German-English Edition*, ed. trans. Gregor & Timmermann (Cambridge University Press, 2011).

only permissible if checked by reason.⁶ The premise of such reasoning is a fundamental obligation that Kant calls the *categorical imperative*, for which he provides three formulations. These formulations dictate the structure of our moral reasoning because they are derived directly from the mere possibility of a rational morality. As the direct formulations of most fundamental moral ‘ought’, they must structure all other obligations, and thus they should be arrived at by any rational agent, rendering them universally valid.

Kant provides three formulations of the categorical imperative:

1. **Universal Law Formula:** “*act only according to that maxim through which you can at the same will that it became a universal law.*” (G 4:421, 6), by which Kant expresses the universality of the moral law;
2. **Kingdom of Ends:** “*So act that you use humanity, in your own person as well as in the person of any other, always at the same time as an end, never merely as a means.*” (G 4:429, 9), by which Kant expresses that the moral law is by and for rational subjects that are equal in their autonomy, and thus that in our actions other rational subjects must be respected equally;
3. **Autonomy and the Good Will:** “*the idea of the will of every rational being as a universally legislating will.*” (G 4:431, 14), by which Kant expresses the necessity of the subject to will the moral law, i.e. to act according to the moral law in one’s autonomy.

From these three principles, Kant argues, all valid moral principles, laws, and rules follow. A formula like the golden rule for example (‘do unto others as you would have them do to you’) can be derived from the first principle, yet is more narrow than it, for it does not involve duties one might have towards oneself and permits those who are better off to ignore potential duties towards the less fortunate (G 4:430, footnote). Additionally, the permissibility of all maxims, which are the descriptions of the intentions that determine our actions, ought to be tested. If I wish to know whether I may, for example, cut in line at the hospital because my stomach hurts particularly bad, I must check this maxim against the categorical imperative and will likely find that it cannot be universalized, because the severity of one’s pain is not a manner in which a queue can be organized.⁷

⁶ This will be discussed in detail in the next section.

⁷ Here we already find much food for discussion, because we must introduce some notion of immediacy when it comes to medical emergencies: clearly someone who is bleeding out should be

From these formulations one can thus derive all the moral duties an agent has, and it is therefore that Kant's moral theory is properly called a deontological theory, i.e. a theory of duties.⁸ Kant distinguishes between perfect and imperfect duties. Perfect duties are those principles according to which we should act because they follow from the categorical imperative straight forwardly. I cannot, for example, will a world in which everyone deceives others when it suits them, because I could never trust someone's word in such a world (which thus runs counter to the first formulation in that such a lie cannot be universalized) (G 4:422, 15). Imperfect duties are those duties which one might desire to be universal (i.e. one desires that everyone act accordingly), but which aren't necessarily so because they do not generate any inconsistencies upon universalization. Kant's example of this is of a man who is wealthy and chooses to disregard the hardships of others, because he contends that all men should fend for themselves (G 4:423, 28). Although this could be a universal law and 'nature could thrive under it', the man cannot will it to be so, for should hardships fall upon him he might require the sympathy or solidarity of others, and thus he would be hypocritical to only reject the concept of charity now that he is fortunate himself.

1.2. Rigorism

By testing the maxims of our actions against the moral law we conclude that they are either obligatory, permissible, or forbidden. More generally speaking, the action is either right or wrong and that rightness or wrongness is inherent to the action. Thus Kant establishes that there are only two ways to relate to the moral law as an agent: one is to acknowledge it and to attempt to check every maxim against it and act

helped before someone who was merely first in line. In the opposite case, where there is for example merely a queue for the post office and a person is bored and wants to cut in line because he finds his boredom unbearable, it seems more intuitive to say we are dealing with 'mere inclination' and this therefore constitutes no argument to cut in line. Yet the severity of one's pain is somewhere in between and only possibly (and not necessarily) constitutes an argument for precedence. Thus deontology is more capable of defending the integrity of a queue as an organizing principle, but a theory such as consequentialism is better at capturing the reasons we might want to breach from that principle in favor of the mitigation of suffering. Deontology would likely be able to respond to this by layering organizing practices, by incorporating clauses of permissible exceptions, etc.

⁸ *Deont-*, from the Greek word for 'duty' or 'that which is binding' and *-ology* meaning 'the science of', Kant was a deontologist in every sense of the word: he studied what it meant to have a moral duty. See G 4:389, 5.

accordingly, and the other is to fail to do so.⁹ Because the moral law is rational, whenever we fail to act accordingly we must be acting irrational, and therefore are instead acting upon some (natural) inclination of self-love. This position is intuitive enough: when e.g. an accountant is asked to falsify financial records for a client, he will reason that this is wrong because his action cannot be universalized (put simply: if all accountants falsified financial records, there would be no point in having accountants) – yet he might be inclined to do so anyway because of greed for money or status, i.e. some form of self-love.¹⁰ Yet this reaches further than that: when we wish to know whether we are permitted to have a third plate of food, we should also check the permissibility against the moral law, because we run the risk of being gluttonous or unhealthy.

It is this binary relation to morality – either we have good or bad will towards the moral law – that leads Kant to embrace the position called rigorism. He does so explicitly in his *Religion within the Boundaries of Pure Reason*, discussing the hypothetical disjunction that man is either morally good or morally bad:

It will readily occur to anyone to ask, however, whether this disjunction is accurate; and whether some might not claim that the human being is by nature neither of the two, others, that he is both at once, that is, good in some parts and evil in others. Experience even seems to confirm this middle position between the two extremes.

It is of great consequences to ethics in general, however, to preclude, so far as possible, anything morally intermediate, either in actions (*adiaphora*) or in human characters; for with any such ambiguity all maxims run the risk of losing their determination and stability.” (R 6:22).

He moves on to defend the rigorist thesis:

⁹ An interesting question is whether this means that every maxim should be checked individually, or whether one may rely on the general moral law. Kant himself writes on the topic that there needn't be a difference between a man who acts on the *letter* of the law or one who acts on its *spirit*, i.e. that there shouldn't be a difference between agents who have a proper understanding of the moral law and therefore can live by it by the letter and agents who conscientiously check their maxims against the moral law in individual situations. More can be said on this topic (for instance, whether such a convergence between a letter and spirit approach is indeed philosophically defensible), but I hope this settles the issue for the present discussion. R 6:30, specifically: “So far as [etc.]”. Cited from Kant, *Religion within the Boundaries of Mere Reason and Other Writings*, ed. trans. Wood & Di Giovanni (Cambridge University Press, 1998).

¹⁰ For Kant's discussion of the concept of self-love, see R 6:27.

On the rigorist's criteria, the answer to the question just posed is based on the morally important observation that freedom of the power of choice has the characteristic, entirely peculiar to it, that it cannot be determined to action through any incentive *except so far as the human being has incorporated it into his maxim* (has made it into a universal rule for himself, according to which he wills to conduct himself); only in this way can an incentive, whatever it may be, coexist with the absolute spontaneity of the power of choice (of freedom). But the moral law is itself an incentive in the judgment of reason, and whoever makes it his maxim is *morally good*. Now, if the law fails nevertheless to determine somebody's free power of choice with respect to an action relating to it, an incentive opposed to it must have influence on the power of choice of the human being in question; and since, by hypothesis, this can only happen because this human being incorporates the incentive (and consequently also the deviation from the moral law), into his maxim (in which case he is an evil human being), it follows that this disposition as regards the moral law is never indifferent (neither good nor bad).

Nor can a human being be morally good in some parts, and at the same time evil in others. For if he is good in one part, he has incorporated the moral law into his maxim. And were he, therefore, to be evil in some other part, since the moral law of compliance with duty in general is a single one and universal, the maxim relating to it would be universal yet particular at the same time: which is contradictory. (R 6:23-25).

This argument follows from Kant's conception of free will as opposed by natural law. Human beings are determined by natural laws, which are causal and universal. Yet as rational beings, humans are the cause of their own behavior and thus autonomous in setting the causal laws for their actions. Kant therefore holds it true that our rational will must be separate from these natural laws, creating a duality between the natural (determined) and the rational (freedom).¹¹ An agent's will is therefore either good because it is rationally determined by the moral law, or it determined by natural causes and therefore bad – within Kant's morality, there simply isn't any space for some third form of neutrality in this regard.¹²

¹¹ A summary of this position can be found in "Kant's Moral Philosophy: 10. Autonomy" in Johnson, "Kant's Moral Philosophy" in *The Stanford Encyclopedia of Philosophy* (Fall 2017 Edition).

¹² Other than perhaps the class of natural inclinations that are checked against the moral law and found permissible, such as satisfying one's hunger, yet these kinds of inclinations are still immoral insofar as they aren't checked against the moral law (leading to e.g. gluttony or over-eating).

There are two ways in which we can conceive of the rigorism Kant defends here. The first is what may be called *rigorism of the will*, which implies that an agent's orientation of the will is either good or bad; good if he wills to act in accord with the moral law, and bad when he acts out of inclination. This binarity seems, as Kant writes, fundamental to deontology, because if one can act partially out of duty and partially out of self-interest, the categorical imperative is no longer categorical. Yet the other manner to conceive of rigorism is *rigorism of the maxim*, which states that there is only one right way to act in a given circumstance, which is according to what the moral law prescribes, and that a failure to act accordingly is to act on a maxim that involves inclination or self-love. It is the latter kind of rigorism that is problematic in the sense that it does not allow for non-ideal circumstances to influence the stringency of the categorical imperative. It is at least not clear which form of rigorism Kant has in mind in the previous citations, for his remarks that "with any such ambiguity all maxims run the risk of losing their determination and stability" and that "the moral law of compliance with duty is a single one and universal, the maxim relating to it would be universal and particular at the same time: which is contradictory" suggest rigorism in maxims: mitigating circumstances require precisely that kind of particularism by defining what might allow for an exception.

Moreover, rigorism of the maxim follows from Kant's equation of the moral law with rationality: the universal law formula is a manner in which we can identify maxims that are inconsistent because they cannot be universalized, thus the formula provides clear, binary answers as to which maxims are permissible and which are not. It is therefore not far fetched to interpret Kant as a full fledged rigorist, and it may well be that this rigorism was on display in his essay *On the Supposed Right to Lie*, which will be discussed in the next section.

1.3. Kant on the supposed right to lie

Before turning to Kant's words on the matter of the permissibility of lying, it must be noted for the purposes of delineating the discussion that the relevant kind of lie is characterized as deception and not as any untruthful declaration.¹³ What is deceptive about the type of lie of interest to the present discussion, is that one makes an untruthful statement that one invites the other to believe that this statement is true.

¹³ Cholbi delineates the discussion similarly. Cholbi, "The Murderer at the Door: What Kant Should Have Said," *Philosophy and Phenomenological Research*, Vol 78., No. 1 (2009): 18. See also Mahon, "Kant on Lies, Candour, and Reticence," *Kantian Review* 7 (1) (2003): 102-133.

Kant (according to Mahon) distinguishes these kinds of deceptive lies from untruthfulness told in contexts of politesse or humor, because in those circumstances the speaker isn't inviting the other party to believe what is said; such untruthful utterances are part of a type of communication in which the truth of one's words (indeed the words spoken) aren't relevant.¹⁴ Yet clearly, in the case of the murderer of the door, the lie we are considering is deceptive: we are misinforming the murderer in an attempt to get him to believe us – in fact, it is very important that the murderer believes us.

Moreover, the case is of particular interest because it would seem that in this situation a deception is warranted: our intuition is that lying to the murderer is either the lesser of two evils (the greater evil being allowing the murderer to proceed unhindered) or not evil at all (that the lie is somehow noble or moral precisely because it prevents a murder). Yet in *On the Supposed Right to Lie from Altruistic Motives*, Kant discusses the possibility of allowing an exception to the principle that one ought to never lie in compromising circumstances, such as this example.¹⁵ In line with his rigorism, Kant simply rejects that this is permissible in that (or any) circumstance:

Thus the definition of a lie as merely an intentional untruthful declaration to another person does not require the additional condition that it must harm another, as jurists think proper in their definition [...]. For a lie always harms another; if not some other particular man, still it harms mankind generally, for it violates the source of law itself. (RL 8:426, 14).

This is a counter-intuitive conclusion to draw, yet several arguments can be derived from Kant's moral philosophy that make it plausible.

When we deliberate upon maxims, we are meant to test them against the formulations of the categorical imperative, and maxims involving deception would, in ordinary circumstances, not withstand such a test. Thus when Kant states that a lie 'harms mankind generally' by 'violating the source of the law itself' it seems he is referring to the notion that deception runs directly counter to the categorical imperative. Certainly, this has been the concern of commentators on the issue, who

¹⁴ Cholbi, "The Murderer at the Door: What Kant should have said" (2009): 109-10.

¹⁵ Translation from Kant, *Critique of Practical Reason and Other Writings in Moral Philosophy*, ed. trans. White Beck (Chicago: Chicago University Press, 1949): 347. RL refers to the German text, "Über ein Vermeintes Recht aus Menschenliebe zu lügen" (1797), found in I. Kant, AA VIII: Abhandlungen nach 1781, followed by the page and line number.

are concerned primarily with the impermissibility of deception under the second formulation of the categorical imperative. Deception seems to run directly counter to this formulation, which states that one should never use other agents as mere means to one's own ends.¹⁶ In this case, the end is to prevent murder and the murderer is being used as a means to reach that end, against his will. It is therefore difficult to find space for an argument that allows one to interfere with the murderer's actions, even if we drastically disagree with his set ends. The deceptive act, after all, would thwart his ends in favor of our own set end to prevent the murder from happening.

Although it seems that deception is particularly problematic under the second formulation of the categorical imperative, the first formulation also raises questions. In line with the universal nature of the moral law, it seems lying 'harms humanity in general' by violating the notion of truthfulness and thus should be rejected in all circumstances.¹⁷ At least it raises another issue for Kantianism, namely the matter of qualifying maxims in order to have them pass the test of universalization. Kantians who want to defend that this maxim can be universalized must qualify it adequately, but this leads to slippery slope problems. Although we may accept that a maxim involving deception to prevent murder may be universalized, what if that murder were part of some institution, such as the military? Or what if the victim were a criminal of some kind? And why is it specifically deception that is permitted, and not the killing of the murderer, which is perhaps equally universalizable? These concerns arise because deception is generally impermissible, and we must therefore find excuses as to why it should be permissible in some particular circumstance.

Finding such an excuse, it turns out, is more difficult than it seems, and is further complicated by the fact that deontology avoids speculation about the consequences of one's actions in favor of deliberating upon maxims in themselves. In light of this, Kant remarks that because one could not know the consequences of a supposedly permissible lie, one could be considered blameworthy if it was the agent's lie to the murderer that resulted in his death.¹⁸ This kind of skepticism also emphasizes what Schapiro considers to be an individualistic conception of agency that follows from the categorical imperative: agents ought to first commit to perfect duties (not to lie) and

¹⁶ Schapiro, "Rigorism and Mitigating Circumstances," *Ethics* 117 (2006): 36.

¹⁷ Note that such an argument is indeed platonic, as if truthfulness itself would require honoring and is not a concept with a function.

¹⁸ Kant, *Critique of Practical Reason and Other Writings in Moral Philosophy* (1949): 347.

consider their imperfect duties (to prevent murder or reduce harm) only after.¹⁹ Arguably this conception of action is expressed in the third formulation of the categorical imperative, because we as agents must commit to the goodness of our will and leave the goodness of the will of the murderer to the murderer.

Finally, returning to the case, it seems that Kant's rigorism does not allow space for dilemma's: an action is either right or wrong, and a dilemmatic situation should absolve once we have sufficient understanding of the categorical imperative.²⁰ Circumstances surely influence the way in which we are permitted to act, but deception of the kind in the case is generally impermissible. The problem posed for any Kantian who wishes to defend the opposite of what Kant defends in *On the Supposed Right to Lie* is that he must argue either (a) that we are allowed to deceive the murderer based on the formulations of the categorical imperative or (b) that the situation is somehow exceptional and that dilemmatic situations do exist. Under the premises of the formulations of the categorical imperative as provided by Kant as well as his rigorism, this has proven difficult, and it is therefore understandable that Kant did not defend this position himself. Kant instead simply argues that deception is not permitted and that interference with the actions of the murderer runs counter to the formulations of the categorical imperative, contrary to our intuitions.

1.4. Kantian apologists

It should come as no surprise that many commentators find this conclusion drawn by Kant problematic. In academic literature, a discussion on this topic emerged around the 1980's, in which several different stances were taken. In this discussion however, the Kantian apologists are more prominent, and historic arguments are supplied to reject the essay. Paton for example objects to a rigoristic reading of Kant in "An Alleged Right to Lie", stating he "[wishes] to estimate the justice of charging [Kant] with rigorism", offering both historic and philosophical arguments. Paton sufficiently nuances his historical arguments to the extent that precedence is given to convincing philosophical arguments. However he does emphasize the questionable circumstances in which the essay *On the Supposed Right to Lie* came about, namely

¹⁹ Schapiro refers to this conception of agency as a community that consists of "an aggregate of self-governors". Schapiro, "Kantianism and Mitigating Circumstances" (2006): 50.

²⁰ Williams uses this term, and I will elaborate on this later. Williams, "Ethical Consistency" in *Problems of the Self. Philosophical Papers 1956-1972* (Cambridge: Cambridge University Press, 1973), 175.

that Kant responded out of combination of injured national pride and old age.²¹ This kind of hermeneutics I find an unfortunate mixture of the history and philosophy: as we might dismiss attribute this essay to Kant's old age, we might attribute earlier essays to his youth and inexperience, and this seems an unfruitful manner of interpreting philosophers and would make sense only in the face of glaring contradictions, which, as has become apparent in the previous, is not the case.

Benton agrees that Paton's historic arguments are insufficient, but does agree with Paton's view that Kant could not have meant what he said in the essay.²² Benton's argument is therefore entirely historical, arguing that Kant did not really mean to be stringent about the particular case of the murderer at the door, but instead wished to argue that Benjamin Constant was being selective in his understanding of the moral law, arguing in favor of an anti-revolutionary movement called the Directory. As was discussed in the previous section, Kant frowned most heavily upon the abuse of the moral law, arguing in *Religion* that even a good act can be made wicked if done out of self-love rather than mere respect for moral law in itself.²³ Thus Benton argues that Kant used the example of the murderer at the door as a wrong example, in which the agent nonetheless acts for the right reasons.²⁴

Yet while this historic argument is more plausible, this again suggests that the problem of rigorism will go away upon rejecting the essay. In the previous sections, I have argued that this was not the case. In fact, Kant was a self-proclaimed rigorist, and although this may not necessarily mean that we shouldn't lie to murderers, it does pose the question when we are allowed to use our best judgment when a situation seems dilemmatic to us. Moreover, it does not seem plausible that Kant would write something he did not mean, especially if such an example were so prone to being misinterpreted. I will therefore turn to the philosophical arguments supplied by Kantian apologists such as Paton and evaluate the philosophical merit of this perspective.

²¹ Paton, "An Alleged Right to Lie" (1954): 193-5.

²² Benton, "Political Expediency and Lying: Kant vs. Benjamin Constant" (1982): 136.

²³ "For whenever incentives other than the law itself (e.g. ambition, self-love in general, yes, even a kindly instinct such as sympathy) are necessary to determine the power of choice to *lawful* actions, it is purely accidental that these actions agree with the law, for the incentives might equally incite its violation." R 6:30-31.

²⁴ Thus in Benton's interpretation Kant ironically abuses the moral law as a rhetorical tool.

1.5. Paton on exceptions to perfect duties

Paton's primary philosophical argument refers to the different types of prescriptions that can be derived from the moral law. He distinguishes (1) moral principles, which are the formulations of the categorical imperative; (2) moral laws, which are derived from the former and admit non-arbitrary exceptions; and (3) moral rules, which are derived from one's private post (such as the duties of a soldier or policeman).²⁵ Paton claims there are no exceptions to moral principles, but there may be exceptions to moral laws when they contradict with one's private duty, in which case a moral rule may override a moral law. It may for example be a moral law to never kill anyone, but it may be one's duty as a soldier to kill the enemy.²⁶ Paton concludes also that the opposite must be the case: "If there may be necessary exceptions to moral laws, then *a fortiori* there may be necessary exceptions to moral rules."²⁷

Paton argues that these necessary exceptions are derived from conflicting duties, and that Kant's distinction between perfect and imperfect duties is indicative of this. We have, for example, an imperfect duty to help others in distress, but it is not clear when, how, or to whom we have this duty and when it is overridden by conflicting (imperfect) duties. Perfect duties on the contrary simply prescribe or prohibit definite action, such as cutting in line and paying our debts.²⁸ Thus imperfect duties do admit exceptions, namely (1) when they are overridden by perfect duties and (2) when they conflict with other imperfect duties. Combined with the argument that some exceptions are not arbitrary but necessary (such as a soldier's right to kill), Paton argues that lying to save a life might not be an arbitrary but a necessary exception, just as for example cancelling a contract for the construction of a dam that would kill thousands, although not truthful, would not be a breach of the moral law.²⁹

Thus Paton argues that, although the moral law is universal, there are non-arbitrary exceptions to it that can be derived from the categorical imperative through proper reasoning. Paton distinguishes between an "ultimate moral principle, which can admit no exceptions, and a particular moral law, such as 'Thou shalt not lie', which

²⁵ Paton, "An Alleged Right to Lie" (1954): 191. Kant makes this distinction in *What is Enlightenment?* For example.

²⁶ Paton states that Kant clearly thought this, but sadly fails to provide a reference to Kant's work. Ibid: 191.

²⁷ Ibid: 192.

²⁸ Ibid.

²⁹ Ibid: 198.

can and does.”³⁰ He draws a parallel to Kant’s theoretical philosophy to illustrate that we must keep an open mind towards the possibility of exceptions to such moral laws:

[...] the principle of causality is on quite a different footing from any particular causal law which may be **discovered by applying the principle**. If we find any breach of a particular causal law, we seek at once to modify it and to discover a more general law in which the principle of causality will not be disturbed, but rather confirmed. Similarly, the moral principle of obeying universal law and never treating rational agents as a means is in no way disturbed if we find it necessary to make exceptions to a particular moral law – such as ‘Thou shalt not lie’ – which is only one application of the moral principle. What we have to show is that the exception, like the particular moral law itself, is also an application of the ultimate moral principle.³¹

Thus Paton has in mind an argument derived from the ultimate moral principle, i.e. the categorical imperative, that proves that the exception that we encounter in practice is non-arbitrary and thus allows for that exception.

Response to Paton

If it is Paton’s aim to qualify Kant’s rigorism, then emphasizing the possibility of exceptions to duties seems a good strategy. It seems plausible to suggest that allowing for exceptions to perfect duties is necessary to mitigate the stringent outcome of the essay. Yet the problem remains that this would require an argument derived from the formulations of the categorical imperative, and this is what is lacking from Paton’s account: we cannot rely on common sense to fill the gap when the formulations of the categorical imperative seems to forbid deception even in this particular circumstance. This is a problem that is general to one principle moral theories, that there must be a clear path of deliberation from the moral principle to the maxim under question. If this line does not exist, intuition trumps theory and the latter becomes mere rationalization.

Paton’s parallel between Kant’s theoretical work and his moral philosophy cannot hold for precisely this reason. Paton seems to have in mind the possibility of arguing that certain exceptions *encountered in moral practice* should warrant us to argue that we have misunderstood the categorical imperative and ought to reconsider why these exceptions might be non-arbitrary. Yet in moral theory, this approach is problematic

³⁰ Paton, “An Alleged Right to Lie” (1954): 197.

³¹ Ibid, my emphasis.

because it advocates for ad hoc arguments when our intuitions conflict with conclusions we draw from the categorical imperative, which is universal. This is the opposite of what Kant had in mind. For Kant, moral agency is dictated by rationality a priori, and our discoveries of counterintuitive conclusions can be no argument to reject our duty to act according to the moral law. The analogy with Kant's theoretical philosophy therefore fails because the latter isn't normative: when we encounter a phenomenon that ostensibly breaches a causal law, we know that ultimately we must come to understand it in terms of causes, because this is how we understand the world.³² Yet if we find ourselves acting contrary to the moral law, we ought to conclude not that we have insufficient grasp of the moral law: on the contrary, because the moral law is normative, we must conclude that we are being immoral and we must adjust our actions.

Paton's response may be that the exception made is non-arbitrary and therefore permissible, and that seems to be precisely his point. However, the precise account of why this exception is non-arbitrary is noticeably lacking. This point cannot stand alone. Conversely, in the previous section we discussed Kant's view on the matter of the murderer at the door, which was perfectly consistent with the categorical imperative. So unless we can provide an overriding account of why this exception is in fact permissible, the argument that it may be non-arbitrary is moot. Additionally, even if these exceptions exist, the duty of honesty seems to follow straightforwardly from the categorical imperative and is therefore a perfect duty, thus a structural argument is needed as to why this perfect duty might admit exceptions in light of some other, conflicting perfect duty.

Insofar as Paton does aim to provide a structural argument, he does so by distinguishing between principles, laws, and rules, hoping to find space for exceptions in this distinction. He claims that a moral rule may result in a non-arbitrary exception to a moral law, such as when a soldier has a private duty to kill, conflicting with the moral law that one shouldn't kill.³³ However he goes on to say that *a fortiori*

³² According to Kant, of course. Quantum mechanics seems to allow for the possibility of understanding the world in terms of *chances*, but I am far too unversed in this topic to do more than merely mentioning it.

³³ This in itself seems to raise issues, because, if there are universal laws that in general prohibit killing, it seems difficult to conceive of a moral society that allows for murder; certainly, an argument must be made as to why this might be a necessary exception. Yet this point in itself may be granted, for there may be ways in which one could argue for such exceptions. Schapiro provides such an argument, and much of Rawls' work on practice rules seems related to this.

this implies that there are also exceptions to moral rules, yet it is not at all clear that this is evident.³⁴ If an agent should conceive of himself both as a moral subject, to whom the moral principles and laws apply, and as a soldier, to whom certain moral rules apply that constitute (non-arbitrary) exceptions to moral laws, yet we should also accept this claim that moral laws sometimes imply exceptions to moral rules, how is this agent to make sense of his duties? Suppose a soldier is ordered to immediately kill an unarmed person, whom he perceives as a civilian, which seems immoral to him. It is perhaps his private duty to obey the orders of his superior and he is allowed to kill in certain circumstances given this private duty, but how should he, as a moral subject, know that this is such a circumstance? None of this is evident, and the moral law, which Kant conceived of as simply universal and accessible via reason (or indeed identical with it), is deluded by such a line of reasoning.

This seems a general problem of the argument, that the universality of the moral law is actively compromised in favor of practical circumstances. The argument is that these exceptions are non-arbitrary, but no convincing argument is delivered as to why this should be so. The clarity and consistency which Kant strives after in his moral philosophy, the one that leads him to embrace rigorism, is here watered down in favor of intuition. Thus this line of defense, that there might be a non-arbitrary defense for circumstances such as the one of the murderer at the door, can only work if such a structural argument is derived from the categorical imperative, or if rigorism can be abandoned in favor of a view that permits mitigating circumstances.

Conclusion of this chapter

In this chapter I have demonstrated that the conclusions drawn by Kant in *On the Supposed Right to Lie*, while counter-intuitive, are not as absurd as they initially seem. The discussion as centered around 'the problem of rigorism' most certainly did not hold for Kant, who was a self-proclaimed rigorist. Kantian apologists have sought to answer this problem mostly by rejecting the essay for historical reasons, but, as has become apparent, there are philosophical reasons for rigorism as well. This is not to say that there is no space for judgment in Kant's moral philosophy, but rigorism is most clearly applicable to those perfect duties that follow straight forwardly from the formulations of the categorical imperative. Therefore either a structural argument

Although these points are not straight-forwardly Kantian, they most certainly could be considered permissible from the perspective of the categorical imperative. More on this in chapter 3.

³⁴ Paton, "An Alleged Right to Lie" (1954): 192.

must be provided that demonstrates clearly why the categorical imperative allows for this instance of deception, or Kantianism must be revised to allow for mitigating circumstances.

Chapter 2: Structural Kantianism

In the previous chapter, it became apparent that Kant's conclusion in *On the Supposed Right to Lie* cannot simply be rejected and that it does follow from his moral philosophy. The next approach would be to provide a structural argument that is perhaps not Kant's own, but that does follow from Kant's formulations of the categorical imperative. Such an approach would suggest that Kant was perhaps wrong in the essay, but only because he failed to see that there was another argument that could be made based on the foundations of his own work.

Cholbi favors such an approach, which I have called structural Kantianism. What he finds lacking in most other accounts of the case of the murderer at the door (including Kant's own) is the argument that we have a duty to protect the victim in the sketched case: "Attempting to establish a permission to lie to the murderer makes it seem as if lying or telling the truth are equally defensible options, but that looks implausible."³⁵ His take on the problem of rigorism in deontology is that maxims are wrong under certain descriptions. We have, for example, a duty not to lie, thus we cannot commit acts that fall under the description of lying.³⁶ However, sometimes actions fall under more than one description and the problem of conflicting duties arises. Therefore, Cholbi is interested in formulating the problem as a conflict of two duties, namely the duty of honesty and some other duty that involves defending the victim. He suggests that a deeper understanding of the latter would serve in understanding why we are not permitted, but in fact obligated to lie to the murderer (or at least to prevent the murder from happening). Then, he hopes to explain why this obligation trumps the obligation to be honest to the murderer in this instance.

2.1. Cholbi on the duty of the preservation of others

Cholbi aims to forward a Kantian duty for the preservation of others by investigating first our duty to preserve ourselves and then exploring the supposedly Kantian notion of a symmetry of duties, which implies that duties that I have towards myself I should also have towards others. He admits that, if his argument would turn out correct, this would conflict with Kant's own words on the matter in *On the Supposed Right to Lie*. He therefore aims to develop a Kantian account that builds upon the core aspects of

³⁵ Cholbi, "The Murderer at the Door: What Kant Should Have Said," *Philosophy and Phenomenological Work*, Vol. 79, No. 1 (2009): 23.

³⁶ Cholbi, "The Constitutive Approach to Kantian Rigorism," *Ethical Theory and Moral Practice*, Vol. 16, No. 3 (2013): 441.

Kant's ethics while rejecting the conclusions drawn in the essay – a position I have called structural Kantianism.³⁷

Kant recognizes a duty of self-preservation which has a rather counter-intuitive but nonetheless consistent foundation. The reasoning is that actions that are not self-preserving, such as self-harm, excessive sexual conduct or drunkenness, and even suicide, are a crime against one's own rational will and autonomously chosen ends in favor of mere inclination. Suicide is conceived of as such an action, because it involves "disposing of oneself as mere means to one's own happiness."³⁸ Self-harm and suicide are thus seen as ways of interfering with one's own autonomy and are therefore forbidden by the duty of self-preservation. This duty aims to preserve an adequate amount of natural capacity towards rational action, and therefore involves notions of moderation and limitation. Cholbi notes however that Kant did not adequately explore the possibility of external sources threatening the preservation of oneself such as in cases of self-defense, and thus Cholbi takes it upon himself to explore those situations as an expansion of the duty of self-preservation.³⁹

When dealing with the duty of self-preservation in cases in which oneself is the origin of the threat, Kant argues that the proper course of action would be removing the empirical conditions that lead to these inclinations (e.g. by seeking help, escaping one's miserable circumstances, buying fewer Snickers bars or bottles of whiskey, etc.). A similar argument can be made for self-defense: when another's aggression or intentions involve the threats towards oneself, we might also have a duty to remove the empirical conditions that threaten our self-preservation.⁴⁰ This would lead to a permissible notion of self-defense: if someone threatens us with a knife, we are at least allowed to disarm them. When someone threatens to strangle us, we might be permitted to detain that person until help arrives, etc. This seems intuitive as well

³⁷ Although this is Cholbi's explicit ambition, I am not sure he succeeds, as he does quite often refer to Kant in a manner that seems to lend the philosopher a special kind of authority. Fortunately, the points derived from Kant can be rejected rather easily, yet unfortunately this somewhat harms the integrity of the argument. Cholbi, "The Murderer at the Door: What Kant Should Have Said" (2009): 21.

³⁸ 'Happiness' should then be viewed in Kant's terms, i.e. as natural inclination. Cited from G 6:423. . Cholbi, "The Murderer at the Door: What Kant Should Have Said" (2009): 24.

³⁹ Ibid: 25.

⁴⁰ Cholbi adds that this is not because of considerations of the happiness of our lives, which may seem odd at first, but is actually more intuitive in the end, for consequentialism must hold such strange notions as that it might be better to allow a murder to happen if it would yield greater happiness. Ibid: 26.

as consistent enough within Kant's theory, for we have a duty to preserve ourselves and we have no duty to assist others in achieving their immoral ends.

Yet it is not evident that the duty of self-preservation is a perfect duty, because Kant recognizes that it might be overridden by other duties (e.g. when a spy must hold information under pressure of torture or even when threatened with death).⁴¹ So the question remains what happens when the duty of self-preservation conflicts with the duty towards honesty, the latter being a perfect duty.⁴²

a. Cholbi on perfect duties

Cholbi therefore turns to the consideration of perfect duties. Rather than addressing the previously posed question, whether self-preservation is a perfect duty, he investigates the stringency of the perfect duty of honesty (or: the duty not to lie). Not unlike Paton, Cholbi considers whether it might be possible that there exists such a thing as a necessary (and thus permissible) lie in Kant's deontology.⁴³ His argument for this possibility is derived from his interpretation of perfect duties. The latter are duties which are derived from a failure in the universalization of certain maxims: when universalizing a maxim is inconceivable or self-defeating, there is a corresponding perfect duty not to act on that maxim. In most cases, lying fails this universalization test because the liar somehow benefits from the lie, but by lying he betrays the truth and his lie will only be believed if the truth is respected. Thus it would be impossible and undesirable if everyone acted on the maxim that involves the lie, which forbids agents from acting on that maxim. As such, interpreters have concluded that lying is never permitted because lies in general fail to be universalized.

Cholbi disagrees with this interpretation, noting that it is not the general descriptor (i.e. lying), but the particular maxim that must be tested against the formula of the

⁴¹ Cholbi uses the term 'absolute' rather than 'perfect.' To avoid confusion, I refrain from following his choice of words. Cholbi, "The Murderer at the Door: What Kant Should Have Said" (2009): 26.

⁴² Cholbi contends that the duty of self-preservation is also a perfect duty, but I am not entirely convinced that this is so, as it contains notions of moderation (e.g. moderation in usage of alcohol and sexual activity), which seem difficult to consider as perfect duties. Yet it would seem that those aspects of that duty that involve the ending of one's rational will can very well be considered perfect duties, and since the case at hand involves murder, we needn't dilute the discussion with the question whether this duty is perfect for now. Ibid: 27.

⁴³ This notion of a 'necessary lie' is apparent lend from Kant's *Lectures on Ethics* ("Notlüge", LE 27:448), but because the source consists of student notes and because the work is rather tentative, Kantians are reluctant to use it as a source. Moreover, the ambition is to provide a structural argument against rigorism, not a historical one. Ibid: 30.

universal law. There may exist a general class of lies that is impermissible, but one must be careful to determine that class precisely, and that determination is reached by universalizing the maxims that build up the class. We are interested in finding the *wrongful* lie and we find that wrongful lie by investigating the particular maxims that contain lying. Cholbi proposes reconceiving of the perfect duty as strict but narrow: when a maxim cannot be universalized, the corresponding perfect duty is binding, but in principle this duty only applies to that maxim or maxims that fall under the same general description of that maxim. Cholbi therefore argues that it is not the case that all maxims involving lies are impermissible because of the corresponding perfect duty towards honesty, but only that there is a general class of wrongful lies built up of maxims involving simple lies out of inclination.⁴⁴

Building on this he notes that, although the lie in the case of the murderer at the door clearly has the intention of deceiving the murderer by manipulating his behavior through deception, it does not do so out of inclination or self-love.⁴⁵ This is similar to the previously established notion that the duty of self-preservation was not a duty based on inclination (i.e. to preserve one's happiness), but was a strictly rational duty. In this regard, deceiving the murderer differs from a typical lie and may therefore not fall under this description.

Although this argument is persuasive, several things remain open. First, what is under consideration is the relationship between perfect duties and the first formulation of the categorical imperative (that of the universal law), not the second formulation (that of the kingdom of ends); in the latter formulation, deception seems generally wrong. Second, we must still carefully analyze the precise relationship between the duty of self-preservation and the duty of honesty: this part of the argument is not yet settled, although space has been made to do so.

Regarding the kingdom of ends formulation, Cholbi states that: "it does not follow from [the prioritization of the value of autonomous rational agency over happiness] that we are obligated to honor another agent's autonomous choices no matter the ends that a given exercise of autonomy is meant to serve."⁴⁶ Kant recognizes a duty to promote the ends of others, but only if those ends are morally permissible. Cholbi inverts this statement to imply also that the autonomy of another agent needn't be

⁴⁴ Cholbi, "The Murderer at the Door: What Kant Should Have Said" (2009): 30-1.

⁴⁵ Ibid: 31.

⁴⁶ Ibid: 32.

honored when they adopt morally impermissible ends.⁴⁷ An impermissible lie, he states, is to manipulate someone away from a morally permissible end through deception, and it is not this kind of lie that is under consideration in the example.

Nor do we have to honor the whole of the set ends of an autonomous agent. He derives this point from Kant's position on retributivism, which recognizes that an agent's usage of their autonomy is important in how they may be treated, but that this does not affect their 'inner moral worth.'⁴⁸ So although inflicting physical pain or restraining someone may be immoral in most cases, the actions of a criminal (i.e. their misuse of their autonomy) might warrant those kinds of responses because of the misuse of his autonomy. If we do not have a duty to assist others in achieving their immoral ends, Cholbi supposes that we may also have a duty to prevent others from achieving their immoral ends. Therefore when deliberating on our own maxims, we may incorporate the immoral ends of others into our maxims, and we needn't consider how we would treat agents in general. This argument is meant to demonstrate that in the kingdom of ends we may still prevent others from committing immoral acts.

b. Deontic symmetry

Having proved that certain lies might be permissible in the formulations of the categorical imperative, Cholbi wishes to demonstrate that we have a duty to preserve others with which the duty of honesty conflicts in the murderer at the door scenario. He argues that there exists a symmetry in the duties we apply to ourselves compared to the duties we apply to others. He calls this the deontic symmetry thesis, which he builds by first introducing two notions:

- *Deontic agents/patients*: a 'deontic agent' is simply one or more agents that perform(s) the act and a patient is the agent(s) acted upon;
- *Deontic valence*: denotes three statuses moral actions may have within deontology, namely *obligatory*, *permissible*, and *prohibited*.

⁴⁷ He does so rather casually, I might add. Cholbi, "The Murderer at the Door: What Kant Should Have Said" (2009): 32.

⁴⁸ Cholbi mentions that autonomy is "not an ahistorical value", but does explore precisely what a historical conception of autonomy would imply, so I will refrain from using this term. Ibid: 33.

The deontic symmetry thesis then reads:

“All other things equal, any act with deontic valence V performed by agent A in which A is also the act’s patient will have the same valence V if another agent B is the patient of A’s act instead, and vice versa.”⁴⁹

Deontic symmetry implies that duties that one has towards oneself (e.g. those that protect one’s own rational autonomy) are special, in that here the symmetry applies. Both formulations of the categorical imperative seem to allow for this. In the realm of ends formulation, we find a kind of impartiality towards whom the recipient of the action is: “whether in your own person or in the person of another, always [treat humanity] at the same time as an end and never as a means.”⁵⁰ Similarly, when it comes to universalization, Cholbi claims it is irrelevant which agent plays the role of agent and which the patient: if I am to treat myself a certain way as a patient, I should treat another similarly.⁵¹

With the duty of self-preservation and the DST in place, Cholbi turns to the case. He first argues that we have a duty to preserve ourselves, as autonomous agents, when threatened with destruction, and thus we have a duty to lie to a murderer that threatened to kill us. Second, he argues that the duty of self-preservation is such a duty to which the DST applies, because there is no argument to discriminate between patients: if it is my duty to preserve my own rational autonomy, it should also be my duty to preserve the autonomy of another.⁵²

c. Refuting objections

Cholbi discusses two objections to this view, the first of which is that lying violates a duty to oneself. He first discusses the rather implausible, literal interpretation of this duty: that we should not deceive ourselves or tell ourselves falsehoods, which, he notes, is both difficult to make sense of as well as irrelevant in this context. The relevant interpretation is the one which states that in lying to others, we violate a duty of truthfulness towards ourselves by treating our morality as a mere means to our natural inclinations. Phrased otherwise, we violate a duty of valuing honesty by

⁴⁹ This quote and all the previous is found in Cholbi, “The Murderer at the Door: What Kant Should Have Said” (2009): 34.

⁵⁰ Cholbi cites G 4:429. Ibid: 36.

⁵¹ Ibid.

⁵² Cholbi discusses two duties (the duty of promotion of the happiness of others and the duty of one’s own moral perfection) that are exempt from the DST because of special arguments, but argues that the duty of self-preservation is not such a duty. Ibid: 38-40.

choosing to say what suits us rather than to speak the truth. Cholbi deflects this by stating that this is certainly an objection when we lie out of cowardice or greed, but this (again) does not apply to every lie, such as those that serve to fulfill another moral duty.⁵³ This argument again hinges on the matter whether lies are wrong generally, or whether impermissible lies are class on their own.

The second objection raises the real concern that lying, in Kant's words, "harms mankind generally, for it violates the source of the law itself." (RL 8:426, 14).⁵⁴ This argument is similar to the one I attributed to Kant in chapter 1, but now comes back as an objection to Cholbi's argument: deception cannot be universalized, for that would render it impossible to trust one another to speak the truth. The simple implication of this argument is that deception cannot be employed as a means to thwart the murderer, because the murderer would have no reason to believe the lie: universalizing the maxim that lying is only permissible as a means to preserve the lives of others renders those means moot, because the murderer would know when we're lying, and therefore it cannot be universalized. "A lying maxim, like a maxim of false promising, depends for its success on a kind of double standard", Cholbi writes. This is to say that it depends on agents to be free riders on the trustworthiness generated by the truthfulness of others.⁵⁵

Cholbi answers this objection rather evasively, pointing out that this it is empirically unlikely that the murderer would know that we would lie to protect the victim, but he is on steadier ground than he seems to think. After all, the murderer would have no reason to ask us where his victim is if he did not believe us to tell the truth – the case hinges on our words counting as deception, otherwise the case becomes trivial. Yet another way in which the objection is then formulated, is that the action cannot be *thought* to be universalized. Perhaps it is possible to be a free-rider of the truth in one circumstance and thwart the murderer, but, the objection goes, the problem for Kantians is precisely that the maxim cannot be employed universally, and thus it cannot hold.

⁵³ Cholbi, "The Murderer at the Door: What Kant Should Have Said" (2009): 41-3.

⁵⁴ Translation from Kant, *Critique of Practical Reason and Other Writings in Moral Philosophy*, ed. trans. White Beck (Chicago: Chicago University Press, 1949): 347.

⁵⁵ Cholbi, "The Murderer at the Door: What Kant Should Have Said" (2009): 44.

This seems the more important objection to address. Cholbi's response is that this is 'a kind of question begging,' because it claims that Kantianism fails to answer how we should ideally act in non-ideal circumstances. Here, Cholbi writes the following:

This is why Korsgaard, Schapiro, and others have claimed Kantians must develop a non-ideal theory in order to make sense of our duties in less than morally perfect circumstances. I have tried to address this particular non-ideal situation, whether we may lie to a murderer, without invoking non-ideal theory, but by appealing to elements already found in Kant's ethical theory. In contrast, this objection digs its heels and effectively ignores the dilemmatic character of the murderer example and the questions it raises about our moral obligations in non-ideal circumstances. Rather than illustrating the wrongfulness of lying to the murderer, the objection instead illustrates the limitations of Kant's universalization test in rendering our obligations sensitive to others' wrongdoing.⁵⁶

This point raises questions about how precisely we should evaluate Cholbi's argument. His stated aim was to answer worries about the case of the murderer at the door 'by providing a Kantian defense of the thesis that lying to the murderer (in the circumstances described in *On a Supposed Right to Lie*) is in fact morally required.'⁵⁷ Yet in the latter quotation, he seems to admit that the case is not settled entirely without invoking non-ideal theory. I will proceed to respond to Cholbi's argument, as well as his treatment of the former objections.

2.2. Response to Cholbi

In summary I take Cholbi's argument to be as follows: (1) the universal law formula permits deception depending on the particular maxim, which should not involve self-love, and (2) the kingdom of ends permits deception when others misuse their autonomy for immoral ends. Therefore, there exists space for a permissible lie within the categorical imperative. (3) In case of conflicting duties, there may exist a *necessary lie*, because it is (a) permissible and (b) obligatory in light of some other duty. (4) Deontic symmetry entails that the duty of self-preservation also applies to others, therefore (5) in the case of the murderer at the door, there is a conflict between the duty of honesty and that of other-preservation. (C) Given 3 and 5, we are obligated to lie to the murderer.

⁵⁶ Cholbi, "The Murderer at the Door: What Kant Should Have Said" (2009): 45.

⁵⁷ Ibid: 19.

The argument as a whole is persuasive, especially if the first two premises are accepted. Yet several of his arguments deserve further scrutiny. Most importantly, Cholbi is not clear on how he wishes to address the non-ideality of this (or any) situation, which is the broader issue at stake. In Kant's rigorism, there is no space for non-ideal theory, and thus a dilemmatic situation arises only as a conflict between inclination and rationality. Cholbi seems to deviate from this perspective, admitting that the situation is dilemmatic or non-ideal. Once this non-ideality is admitted, we may put the formulations of the categorical imperative into a new perspective, as Cholbi does. Yet it is not clear that the categorical imperative permits deception of any kind in the manner conceived by Kant. In Cholbi's conception, there is a strong reliance on judging both the intentions of others as well as the consequences of actions to deviate from Kant's rigorism – such qualifications of the categorical imperative only make sense once non-ideality has entered into our deliberation. Finally, the deontic symmetry thesis is not unproblematic because of the subjective nature of Kant's deontology, in which the autonomy of agents decides whether an act is right or wrong. I will elaborate on these points in the following section.

a. Cholbi and non-ideality

At the end of the article, Cholbi admits that the situation is non-ideal, while in the preceding he has argued that we have an obligation to lie to the murderer. What is problematic about his approach is that he lends from the dilemmatic or non-ideal aspects of the dilemma to demonstrate that the dilemma can be absolved after all, and that we are obligated to lie. It seems to me he cannot do both. We may interpret Cholbi to mean either one of two things, the first being that the dilemma is absolved and that Kant was right for being a rigorist, because the dilemma wasn't a dilemma after all; the second being that Cholbi admits that non-idealism must be admitted into the categorical imperative, and (perhaps) that the formulations thereof are robust enough to deal with such non-idealities. I will elaborate on this in the following

What is meant by the dilemma being absolved, is that for Kant, there is only one solution to any dilemma, precisely because in every circumstance there is only one right way to act, namely conform the categorical imperative.⁵⁸ This right action must therefore be derived from the formulations of the categorical imperative. However, if we allow the misconduct of others into our maxims, we become open to dilemmatic situations after all. There will be things we cannot know of a situation, intentions of

⁵⁸ Williams, "Ethical Consistency" (1973), 175.

others must be incorporated into our maxims even though we cannot verify them or their history, etc. Moreover, it will not be clear how we, as agents, ought to act ourselves: a lie may be considered permissible when facing a murderer, but is murder permissible? If it is no longer clear how we ought to act based on the formulations of the categorical imperative, then the stringency of the moral law is mitigated. In this regard, the actions of others cannot mitigate the stringency of our own duties.⁵⁹

The problem that arises is that Cholbi seems to think he has not invoked non-ideal theory, when in fact he has.⁶⁰ By incorporating the immoral intentions of others into our maxims and by asking whether we are allowed to commit immoral actions to thwart others from behaving immorally, Cholbi admits to the dilemmatic character of the situation. Kant's argument in *On the Supposed Right to Lie* avoids doing precisely that: in Cholbi's words, it assesses maxims by asking if they could govern our interactions in a world of full compliance, and then instructs agents to act accordingly. If we interpret Cholbi's approach as loyal to Kant's core tenets, thus arguing that we need no 'lesser of two evils' or non-ideal understanding of the dilemma, we might interpret him to argue that the dilemma is absolved by him having proved that we have an obligation to lie to the murderer, and that he has therefore successfully defended Kant's rigorism while simultaneously rejecting the need for non-ideal theory.⁶¹ Yet the manner in which he comes to that conclusion, that we have an obligation to lie to the murderer, is by incorporating the immoral intentions of the murderer into our maxim before testing it against the universal law and kingdom of ends formulations. Thus a problem arises: Cholbi requires the non-ideal elements of the case to argue that we are obligated to lie to the murderer in ideal theory.

When we reconsider his arguments derived from the formulations of the categorical imperative, their problematic aspects become more apparent. His argument that the particular maxim should be considered when tested against the universal law formula and that a lie to a murderer is universalizable, is problematic because it

⁵⁹ Schapiro describes the problem similarly: "To think that others' wrongdoings can in any way mitigate the stringency with which the moral law applies to us is, presumably, to deny its categorical status." Schapiro, "Compliance, Complicity, and the Nature of Nonideal Circumstances," *The Journal of Philosophy*, Vol. 100, No. 7 (2003): 330.

⁶⁰ Cholbi, "The Murderer at the Door: What Kant Should Have Said" (2009): 45.

⁶¹ This seems in line with his ambitions to "[provide] a Kantian defense of the thesis that lying to the murderer (in the circumstances described in *On a Supposed Right to Lie*) is in fact morally required." Yet it seems at odds with his later remarks on the "dilemmatic character" of the situation. *Ibid*: 19 & 45.

incorporates the intention of the murderer as a justification of the lie. Similarly, we could justify a particular maxim that involves killing or torturing the murderer and find it universalizable: after all, murderers are non-ideal, and if (a) we abandon the ideals set in moral deliberation but (b) do so without clear guidance, then the immoral actions of others can make most of our immoral actions justifiable.

Similarly his argument for the permissible lie in the kingdom of ends formulation, that the immoral action of the murderer permits us to deceive him, is problematic when it is considered that we must deceive the murderer, thus committing to an immoral act ourselves. On that topic, Cholbi writes: “Kant claims that we are obligated to promote others’ ends (i.e., others happiness) only if those ends are themselves morally permissible. I take that to be an indirect statement of the notion that whether we must honor another’s autonomy on a given occasion may depend on whether that autonomy is being exercised in morally permissible ways.”⁶² Yet this inversion is not at all clear in the case in which we thwart the murderer through immoral actions: not promoting immoral ends is dissimilar from actively thwarting the actions of others through immoral actions.

Although Cholbi’s argument does provide an account of how the formulations of the categorical imperative might still be action guiding, albeit in a less stringent way, once rigorism has been abandoned, the questions raised in chapter 1 can be raised here again: Why can precisely deception be universalized when we are faced with a murderer? How should we act given more information, such as when the victim has also acted immorally, or how should we act given a lack of such information? Moreover the underlying question, how conclusions drawn in ideal theory (i.e. that we shouldn’t lie) affect our actions in non-ideal situations, remains unanswered.

b. Cholbi’s deontic symmetry thesis

There is a clue, it seems to me, in Cholbi’s deontic symmetry thesis as to where his views actually differ from Kant in a relevant way. The deontic symmetry thesis interprets the universal law formula and kingdom of ends formulation as indifferent towards who the agent and who the patient is. Yet from the perspective of Kant’s moral theory it is not evident that this is the case, especially when the third formulation of the categorical imperative is considered. For Kant, it seems more important that the individual agent, as a moral subject, employs his autonomy with

⁶² Cholbi, “The Murderer at the Door: What Kant Should Have Said” (2009): 32.

respect for the categorical imperative. Phrased otherwise, the goodness of the agent depends on the orientation of the will of the agent towards the moral law. This conception of agency centers around the moral subject, and from this perspective the agent and patient of a duty are not as exchangeable as Cholbi suggests.

This I think explains Cholbi's trouble making sense of Kant's notion that the duty of honesty is one we have towards ourselves and mankind generally, rather than to other agents. To Kant, we have this duty towards ourselves because being deceptive betrays the orientation of our will, in which we are allowing natural inclination knowing that deception is neither universalizable, nor justifiable in the kingdom of ends formulation. In the latter sense, it harms mankind generally, namely for arguments similar to Cholbi's "double standard" argument: humanity depends on honesty in interactions, and thus dishonesty is harmful to humanity.

In Kant's conception, perfect duties are therefore duties that involve clear rational inconsistencies that ought not be violated by moral subjects. The duty of honesty or the duty of self-preservation are clear examples of this, because we know precisely when we are lying or when we are damaging our capacity for rational agency.⁶³ It is therefore not necessarily plausible that the deontic symmetry thesis would yield perfect duties in Kant's conception, because these duties are often not rationally inconsistent in the manner in which the duty of honesty is.

What the deontic symmetry thesis presupposes is a move towards a different conception of agency, namely one in the conscientiousness of one's action depends on the ability to make one's will public. The universalization of one's actions then depends more on the fairness thereof, i.e. the interchangeability of myself with another agent having no impact on my choice of action. The kingdom of ends functions in a similar fashion in this conception: your ability to pursue ends is just as important as mine, and my actions can be justified in light of this. In contrast to Kant's conception, the goodness of my will is less dependent on my subjective orientation of the will, but more so on my ability to justify my actions to others or to make my will public.

⁶³ More or less: there are of course plenty of cases in which it is not clear whether we indulging in pleasures that are harmful to our rational autonomy, but the duty is perfect at least in the cases in which we are certain of this, such as in cases of suicide, addiction, etc.

Conclusion of this chapter

Cholbi's account exposes some of the problems that rigorism has when dealing with non-ideal situations. Unfortunately, his arguments regarding the formulations of the categorical imperative already presuppose the incorporation of non-ideality into our maxims, which begs the question how the conclusions drawn in ideal theory should relate to dilemmatic practices. Cholbi, in my view, fails to answer this question, because he does not answer it explicitly: it son the one hand explicitly his ambition to respond to the case from Kantian ideal theory, yet he does so by incorporating precisely those aspects that make the case dilemmatic. In the next chapter, I will discuss Schapiro's revision of Kantianism. Her project is to address the relationship between ideal and non-ideal theory explicitly.

Chapter 3: Schapiro's revised Kantianism

Schapiro discusses Kantian rigorism in her article *Kantian Rigorism and Mitigating Circumstances*.⁶⁴ She derives the essence of the problem of rigorism from what she calls the *deontological thesis*, which is that an action is intrinsically right or wrong, based on arguments derived from the categorical imperative: if actions are conceived of as intrinsically right or wrong, there seems to be no space for mitigating circumstances.⁶⁵ In the article, Schapiro focusses on the second formulation of the categorical imperative in particular, i.e. that one should treat another agent as an end and not as mere means. From this she derives what she calls the standard argument against (paternalistic) deception, which is that not providing the information another agent needs to act in accord with her goals violates the autonomy of that agent. This kind of violation of autonomy she calls “deception-as-interference.”⁶⁶

3.1. Problems of the standard argument

Schapiro investigates the possibility of mitigating circumstances in common situations, such as when we tell lies to children or the mentally limited. Insofar as these situations are morally problematic, they are so under the second formulation of the categorical imperative, i.e. as violations of autonomy: we ought to let other agents set their own ends and respect those ends. Yet in these common situations, she distinguishes two ways of justifying this interference, namely (a) because they are persons who are likely to set bad ends (e.g. children or the mentally limited), and (b) because they are persons who are (sometimes) incapable of making choices that *count as their own* and therefore don't (always) have autonomy.⁶⁷ Option (a) is not open to the Kantian, because it is precisely such a violation of the autonomy of another person

⁶⁴ Schapiro, “Rigorism and Mitigating Circumstances” (2006): 32-57.

⁶⁵ *Ibid.*, 35.

⁶⁶ *Ibid.*, 38.

⁶⁷ I am not sure if it is philosophically defensible to say that agents are sometimes autonomous and sometimes not, as it would be difficult to estimate when the agent is autonomous from one situation to the other. Certainly, there would always be a kind of paternalism over the person: an addict might give up his autonomy when it comes to his heroin usage and prefer someone else make all decisions for him regarding his treatment – likewise, someone who raises a child or takes care of someone mentally limited might think the other person capable of autonomy when it comes to one situation, but not to another. Yet this judgment in itself is made by the caretaker, rendering it ultimately a case of paternalism. This is beside the point of this paper, but interesting nonetheless. The easiest answer is to ignore my parenthesis and simply say that an agent is not sometimes autonomous, but either is or is not.

which is forbidden by the second formulation of the categorical imperative, yet she argues that the second option (b) can be justified. Because the agents aren't autonomous and therefore aren't agents, they cannot be deceived at all. Instead, we are performing the act of "intentional falsehood telling."⁶⁸ Thus when we steer the actions of infants by telling half-truths we are not actually deceiving them, we are merely telling falsehoods – in fact, we can't deceive them because they lack autonomy.

Yet, Schapiro notes, distinguishing honesty and deception from intentional truth and falsehood telling leads to new problems. First, conceiving them as two separate categories of action is unintuitive and therefore problematic, because it would seem the two categories are intimately related and that their moralities inform one another. We might for example tell a half-truth to spare the feelings of an infant, but we wouldn't consider it right to outright deceive them (at least not on relevant topics), and the reasons not to do so seem similar to the reasons we have for not deceiving an autonomous adult.⁶⁹ Second, it is precisely the point of certain marginal cases that the circumstances excuse the deception, i.e. 'that circumstances C forced us to do act A in spite of standard S.' Thus it seems we would want to distinguish what is 'the right thing to do' from what is 'the lesser of two evils, given the circumstances.' In doing so, we would recognize that there is a class of actions that is not straight-forwardly justified via moral reasoning, but where circumstances regrettably make deception the most choice-worthy course of action, and thus excusable.⁷⁰

Schapiro derives a third problem from the standard argument for paternalistic deception by introducing the marginal case of the murderer at the door, which she calls a case of "defensive deception."⁷¹ She concludes that the standard argument does not permit this: the murderer's autonomy is not impaired and therefore the argument above doesn't hold, and because no space is made as of yet for a kind of excuse in light of mitigating circumstances, the marginal case remains problematic.

Schapiro therefore argues that deontology would require a type of *prospective excuse*, which is to say that an agent may deviate from the standard in certain excusable circumstances that can be established beforehand, i.e. when circumstances are such

⁶⁸ Schapiro, "Rigorism and Mitigating Circumstances" (2006): 38-40.

⁶⁹ Perhaps here it is relevant that a child is potentially autonomous, and thus that we are preparing them for autonomy, but this argument cannot hold, because we wouldn't consider it right to deceive mentally limited persons either.

⁷⁰ Ibid: 40-3.

⁷¹ Ibid: 43.

that deviation is excused. Such an argument would avoid being ad hoc or after the fact, but would nonetheless excuse agents from choosing the lesser of two evils in mitigating circumstances. This, she argues, would fix all three the problems, and she continues to develop a revision of the standard argument that allows for such an excuse.⁷²

3.2. Schapiro on honesty: letter and spirit

Schapiro hopes to solve all three problems with the standard argument with her own, revised Kantianism. To recap, these problems were: (1) that the separation between honesty and intentional truth telling is unintuitive and problematic; (2) that the marginal cases don't seem to justify certain actions, but rather seem to excuse them; and (3) that it doesn't resolve the truly problematic marginal cases such as the murderer at the door.

She does so by first noting that if deception is wrong because it interferes with the autonomy of others, honesty must be right because it respects the autonomy of others. Thus we shouldn't ask when deception is permitted, but when a deviation from the categorically binding standard of honesty is permitted.⁷³ One such situation we find when the agent we are interacting with is not (fully) autonomous.⁷⁴ In Schapiro's conception, this is not a wholly separate class of action that is completely amoral, as is suggested by the standard argument, but is instead a depraved form of honesty. (Indeed, although Schapiro does not discuss this, one could argue that under the third formulation of the categorical imperative, that of the good will, agents should still have the will to be honest, even if they can't be.) Schapiro therefore separates the spirit of honesty, which must always be lived up to, from the letter of honesty, which are merely the actual words spoken.⁷⁵ Actions involving intentional truth telling are in this conception a form of honesty that is depraved because the agent that is spoken to is not autonomous.

⁷² Schapiro, "Rigorism and Mitigating Circumstances" (2006): 41 & 45. Although she does not mention the term 'a priori' explicitly, it would seem to me that this is what is implied by her description of the excuse as 'prospective.' After all, we cannot speak of a deontological argument if it is not a priori, i.e. if its wrongness or rightness is not concluded from the argument itself. The argument for the excuse would therefore have to follow from an analysis of the action itself.

⁷³ Ibid: 45.

⁷⁴ We might say that we can only be as honest with others if they are autonomous.

⁷⁵ Similar arguments can be made as to why we should not want to harm animals (as non-autonomous entities), because despite not necessarily being immoral it might damage our good will.

What does this devaluation mean precisely? To Schapiro, it refers to what she calls the *background constitutive context*. Under the surface level of our actions we make assumptions about the autonomy of the agents around us, and although our behavior should always be governed by the same moral principles (such as the one that states that deception is wrong), sometimes it seems we have no choice to deceive because background conditions are less than favorable. Specifically, these conditions are corrupt when the agents we interact with cannot or do not respect the reciprocal relationship necessary for moral interactions. Schapiro thus hopes to create space within deontology for a prospective excuse: when background conditions are not in order, e.g. when the autonomy of the individuals we interact with is lacking, we are excused from choosing not honesty both in letter and in spirit, but honesty merely in spirit, in which case the letter of honesty requires improvisation.⁷⁶ Conversely, paternalistic deception may in those circumstances be permissible in letter, but it is never permissible in spirit.⁷⁷

As a result of this separation, Schapiro has fixed the first two problems of the standard argument: (1) intentional-truth-and-falsehood-telling are no longer separate *types* of action, but rather are different *forms* of honesty and deception; and (2) space is made for the concept of a preemptive excuse, namely when background conditions are corrupted. The latter implies that there may now be an adequate and intuitive conception available for a 'lesser of two evils' scenario.

Thus Schapiro turns to the third problem of the standard argument, which involves the case of the murderer at the door. But because the murderer doesn't lack autonomy, it no longer makes sense to justify deception-as-interference, because his autonomy renders interference impermissible. How can a reference to the background conditions then excuse deception? Schapiro claims we must conceive of the deception not as an interference, but as a *refusal*. This requires a shift in the way we conceive of action and deception: "Whereas deception-as-interference presupposes a conception of the moral community as an aggregate of self-governors bound by obligations of noninterference, deception-as-refusal presupposes a conception of the moral community as a realm of joint participation in a shared activity."⁷⁸ Honesty, she

⁷⁶ Schapiro, "Rigorism and Mitigating Circumstances" (2006): 47.

⁷⁷ There are also circumstances in which one can be honest in letter, but not in spirit, e.g. when a criminal happens to benefit from telling the truth, but would have happily lied if it would have suited him better. Ibid: 48.

⁷⁸ Ibid: 50.

claims, should be seen as form of interaction and not merely a form of action. The spirit of honesty then requires an addition, which is that it implies a common, co-legislative standpoint. The murderer at the door clearly does not embrace that standpoint when he wishes to kill another agent, and thus honesty, as a reciprocal interaction, becomes impossible with him.

The intention of the murderer is not a justification for deception, because this would run counter to the second formulation of the categorical imperative. Yet it may excuse our actions, because deceiving him in a circumstance in which he intends to murder is preemptively permitted because honesty has become impossible. After all, the murderer has removed himself from the moral community by adopting ends that are at odds with the latter.⁷⁹ Schapiro considers this a revision of Kant's individualistic conception of action, in which the rightness of an action is limited to the agent's own intentions. As was discussed previously, the honesty of an agent depends primarily on the orientation of his will. Yet in Schapiro's view, honesty is an interaction which becomes impossible when the other party disrespects the normative premises of moral interaction altogether. In this view the second formulation of the categorical imperative, the kingdom of ends, should be conceived of as a moral community that forms the foundation of autonomous moral action. By intending to kill, the murderer intends to violate the integrity of the moral bedrock of all moral action.

Conceived as such, honesty and deception are not modes of action, but of interaction, "defined against a normatively structured background relationship of reciprocity between autonomous agents."⁸⁰ When this relationship is corrupted, honesty and deception become impossible in their full sense, and "constrained forms of deception can become permissible."⁸¹ This also separates the ideal from the non-ideal, in that ideally these background conditions are in order and we can deliberate upon our actions as though they were, yet we must deviate from those actions when they are corrupted. Thus Schapiro hopes to have responded to all three the complaints, marrying our intuition that we may deceive the murderer at the door with deontology by developing an account of a preemptive excuse and shifting the discussion from individual to shared agency.

⁷⁹ Schapiro, "Rigorism and Mitigating Circumstances" (2006): 52.

⁸⁰ Ibid: 56.

⁸¹ Ibid.

3.3. Cholbi's 'anything goes' objection

In a recent article, *The Constitutive Approach to Kantian Rigorism*, Cholbi criticizes the former approach. Cholbi's argument centers around the problem that occurs when we were to expel someone from the moral community. In her constitutive approach, Schapiro argues that murderer's action is not merely a moral offense, but is in fact a refusal or a betrayal of the moral standpoint altogether.⁸² This notion of a betrayal rather than an offense is both necessary and intuitive: an offense to the standards of reciprocity is something that could occur quite commonly, e.g. when we fail to recognize our duties correctly or adequately.⁸³ It would be a difficulty of the constitutive approach if these kinds of 'offenses' would warrant an excuse of some kind for other agents to deviate from their duties towards that person. But the difference is that the murderer intentionally betrays the underlying normative framework of reciprocity and it is in this sense that his intentions are relevant.⁸⁴ A mere offense is a misinterpretation or misapplication of the moral law, but not a failure to respect that underlying framework.

Schapiro also wants to uphold the Kantian notion that this betrayal does not cause the murderer to be expelled from the moral community entirely: he has not 'exited' the moral community altogether and we are not allowed to treat him however we wish. Yet, as noted, she simultaneously holds that we are excused to deviate from our otherwise categorical duty to be honest by deceiving the murderer precisely because of his betrayal. Honesty, after all, is impossible with agents who fail to recognize that others are deserving of equal and reciprocal treatment.⁸⁵ Thus Cholbi raises the 'anything-goes' objection: it is unclear why it is precisely *deception* that is excused as a defensive action, and why we should not e.g. kill or torture the murderer, i.e. why the murderer isn't considered outlawed. After all, it is not merely or precisely the spirit of honesty that is being betrayed by the murderer (which might count as a mere

⁸² Cholbi, "The Constitutive Approach to Kantian Rigorism," *Ethical Theory and Moral Practice*, Vol. 16, No. 3 (2013): 443.

⁸³ Examples of this are plentiful, such as when an agent in financial troubles receives an unwarranted check from the bank and decides to keep it. Assuming an argument can be made that this is immoral (which I won't make here), such an agent might think that the enormous wealth of large companies makes her action universalizable and that it is in fact okay, when really her judgment is affected by her financial troubles.

⁸⁴ Schapiro refers to this (I think) when she describes the relevance of the orientation of the will of the murderer. Schapiro, "Rigorism and Mitigating Circumstances" (2006): 52.

⁸⁵ Cholbi, "The Constitutive Approach to Kantian Rigorism" (2013): 444.

moral offence), but the underlying normative framework that constitutes all forms of moral interaction.

One rather obvious response would be that something must be said about the severity of the intervention. Dishonesty is clearly a more benign response than killing the murderer, and Cholbi states that deontology has responded to similar objections by introducing norms of proportionality.⁸⁶ A similar strategy seems open to Schapiro by arguing that agents should still want to live up to the various ‘moral spirits’, of which honesty is only one and non-violence might be another, and in aiming to live up maximally to each of them deceiving the murderer is simply the best option.⁸⁷ Yet Cholbi asserts that the constitutive approach cannot rely on deontological arguments of proportionality, because this argument is again justified only in light of the normative framework that the murderer has already betrayed. So even if an argument can be developed why it is more morally appropriate not to resort to violence but merely to deceive someone, that argument would still not apply, because the murderer has betrayed precisely the normative framework that makes such an argument applicable.⁸⁸ It is therefore a problem, according to Cholbi, that Schapiro wishes to excuse precisely deception in light of the murderer’s betrayal of the normative framework, because the latter is a betrayal of the entirety of morality: if there is any obligation to save the victim, then we may use any means to do so, because we have no duties towards the murderer whatsoever. According to Cholbi, Schapiro’s argument therefore makes any reaction, no matter how immoral, permissible.

Conclusion of this chapter

In the former, Schapiro’s response to the murderer at the door case has been discussed. Schapiro proposes a shift in the way the categorical imperative is perceived, emphasizing its relational aspect. In this conception, a moral rule functions to constitute moral interactions that become corruptible when the normative background conditions are not in place. The latter is the case when the murderer asks for information that will enable him to kill: he then betrays the moral

⁸⁶ After all, the standard argument would also require such an argument, unless it wants to hold that it truly doesn’t matter what we tell the non-autonomous, such as infants. Cholbi, “The Constitutive Approach to Kantian Rigorism” (2013): 445.

⁸⁷ Cholbi mentions this deontological solution of living up maximally to various duties and responding proportionally. I can’t develop the full account here, but it seems plausible enough. Ibid: 445.

⁸⁸ Ibid.

community at large. Yet Cholbi objects that, with this betrayal, the murderer has effectively removed himself from moral consideration altogether and thus we may treat him in whatever way we like. Whatever norms of proportionality we might otherwise employ are lost because the normative framework that validates such norms are also betrayed by the murderer: in violating the law, he has outlawed himself.

Is this objection valid? A response to it must express how the agent relates to the moral law or normative framework once another agent, e.g. the murderer, has betrayed it. If we enter into a realm of surrogate morality where we are allowed to improvise with the moral law, then Schapiro's conception of Kantianism must somehow explicate the relationship the agent still has to the ideal of the moral law. I believe Schapiro's conception can be understood in a manner that makes sense of this relationship, which is really the relationship between the ideal of morality in an empirically non-ideal situation. The key to understanding this conception lies in her discussion of deontology in social practices and the manner in which participants of a practice that has been corrupted or usurped by a unilateral will.

Chapter 4: In Defense of Revised Kantianism

The anything-goes objection as raised by Cholbi is aimed at the constitutive approach, but can be broadened to incorporate any Kantian moral theory that allows for non-ideal particularisms to enter into ones maxims. Thus ironically, this objection can not only be leveled at Schapiro, but applies to Cholbi as well when he wishes to consider for example particular maxims regarding permissible lies. If we for example ask whether we are allowed to deceive a murderer, we may as well ask whether we are allowed to kill a murderer, and we will find the action is universalizable precisely because the murderer acts immoral and, from the ideal perspective, shouldn't employ his autonomy in this manner in the first place. Kant's avoidance of the non-ideal and his rigorism gains appeal in light of these issues. If there is such a thing as categorically binding moral law that dictates the principles that would govern an ideal world and agents have access to it via their rationality, would full compliance to its norms not be the only sensible position to take in moral theory?

The discussion about the possibility of non-ideal theory and the question of full versus partial compliance was introduced by Rawls and has been prominent primarily in political philosophy.⁸⁹ In this domain, questions arise regarding how much agents ought to do when others don't do their part, or when they know their actions will have little to no influence (or even negative consequences) even though they seem like the right thing to do (such as going to a protest against some political injustice, even if one is the only protester and is likely to get in trouble for it).⁹⁰ In such matters, it is obvious that under ideal circumstances one should do X, but given the non-ideal circumstances, it might be better not to do X.

What is of interest for the present the discussion is the question whether a non-ideal domain of reasoning can exist for the individual agent and, if so, when we enter into it. The only viable alternative seems to be that Kant's moral theory simply entails rigorism, a position in which non-ideal situations will disappear once we understand our duties sufficiently, and the formulations of the categorical imperative are the means to that understanding. In *Ethical Consistency*, Williams describes this response of ethical theories to dilemma's as an attempt to absolve the dilemma by providing an answer, which, in his view, is not the form that a real dilemma should

⁸⁹ Valentini, "Ideal vs. Non-Ideal Theory: A Conceptual Map," *Philosophy Compass*, 7 (9) (2012): 2. See also Rawls, *A Theory of Justice* (Cambridge, Massachusetts: Belknap Press of Harvard University Press, 1971).

⁹⁰ Valentini, "Ideal vs. Non-Ideal Theory: A Conceptual Map" (2012): 3-4.

have. The point of a moral dilemma is precisely that it could give agents moral regret for having to have chosen the lesser of two evils (even if the agent agrees with her choice retrospectively): “Moral conflicts are neither systematically avoidable, nor all soluble without remainder”, he writes.⁹¹ Williams calls such a position a form of ‘glib moralism,’ because it denies the existence of dilemma’s in favor of theoretical consistency.⁹² Kantian rigorism in the full sense can be accused of doing precisely that.

This issue lies at the heart of this thesis, because it enquires directly into the manner in which agents relate to the ideal of the moral law in non-ideal circumstances. For Cholbi, allowing a non-ideal realm of deliberation to exist is similar to admitting that in certain circumstances, our actions cannot be justified and therefore ‘anything goes.’ This seems similar to Kant’s remarks on rigorism, that “with any such ambiguity all maxims run the risk of losing their determination and stability.” (R 6:22).

Thus the question becomes whether Schapiro’s Kantianism is equipped to deal with this complaint. In another article about social practices, she discusses similar tensions within the realm of social practices. Although she does not apply this conception to morality at large, she does end this article stating that it may be possible to conceive of morality as an omniscient practice that underlies all moral interactions:

[...] what if there were a non-optional practice, a role we had to occupy on pain of not being agents at all, and a set of rules we had to follow on pain of not doing anything at all? If abandoning agency is not humanly possible, then those rules would be in the fullest sense categorical imperatives, and we would be bound to adhere to them under all conditions. But if those rules were categorically binding in virtue of their constitutive role – in virtue of their fitness to constitute a shared will and a shared form of activity attributable to that will – then they would by the same token be vulnerable to corruption; in principle, there would be some threshold below which noncompliance could force the letter and the spirit of the moral itself to come apart. In this unfortunate circumstance, the conscientious agent would be forced to ask

⁹¹ Williams, “Ethical Consistency” (1973), 179. See also 175 for his criticism on absolving the conflict.

⁹² Williams, “Ethical Consistency” (1973), 173. Schapiro, “Compliance, Complicity, and the Nature of Nonideal Circumstances” (2003): 334.

whether his *regard for the universal form* of his maxim had itself been rendered something of a *mere formality*.⁹³

In what follows, I will discuss Schapiro's conception of social practices, aiming to finally see if this can enrich our understanding of constitutive Kantianism so as to respond to the anything-goes objection.

4.1. Social practices and non-ideal conditions

In her article "Compliance, Complicity, and the Nature of Nonideal Circumstances," Schapiro lays the foundation for what she calls "two-level Kantianism."⁹⁴ In this article, Schapiro aims to develop a response to Kantian purism by finding "an ideal/nonideal distinction of the right kind". She does so by forwarding a particular conception of social practices as rule governed activities that depend constitutively on the compliance of others with the practice. Her aim is to find the dilemmatic tension within the practice, and not by raising concerns of the efficacy of the practice to realize some goal in consequentialist terms.⁹⁵

First, she investigates situations in which we might find that agents are 'making the moves' demanded by the rules of the practice, but are doing so for reasons external to the practice. Schapiro's example of this is the case of a sham negotiation, such as when you try to negotiate properly by aiming to come to a mutually acceptable agreement, whereas I am merely stalling the negotiation to buy time.⁹⁶ I am then merely making the moves of the practice without being committed to the practice end, whereas you are. Clearly, my form of negotiating is a sham. Yet how should we describe your behavior? Are you still negotiating, when I am merely buying time? "If it makes sense to raise this question," Schapiro writes, "this suggests that we take there to be some threshold beyond which the noncompliance of another can have bearing on the integrity, and not merely the efficiency, of your rule-governed action."⁹⁷

Schapiro rejects the conception of a practice of which the constituency depends entirely on its efficacy, for this would render a practice a mere *method* to achieve a certain end, meaning that all practices are best analyzed in terms of their

⁹³ Schapiro, "Compliance, Complicity, and the Nature of Nonideal Circumstances" (2003): 355.

⁹⁴ Ibid: 336.

⁹⁵ Ibid: 332-4.

⁹⁶ Ibid.

⁹⁷ Ibid: 337.

consequences.⁹⁸ She wants to avoid this conception firstly because, were a practice a mere method, that would imply that we are justified in abandoning it whenever we deem abandoning more efficient. This, it seems, is not how practices function: they have an internal normativity as a practice, and not merely relative to some set end. If for example my participation of the legal system is a method to achieve justice, I could abandon it whenever it failed to do so, for example when a criminal walks free due to a lack of evidence. Schapiro's point is that a practice is a procedure of "making a problem, along with its solution, count as *ours*."⁹⁹ The reason I have to care about the legal system being just is because it counts for all of us: the insufficient evidence that allowed a criminal to walk free is the result of principles that exist to ensure justice for all its participants and allows us to agree on a system that constitutes such justice. In the latter sense, practices are a separate source of agency and actions within them are public actions constituted by shared practice rules. "In a practice," Schapiro writes, "actions are attributable to a shared will because and insofar as participants make reciprocally binding claims upon one another to adhere to the rules of the practice."¹⁰⁰ Here again we see Schapiro's position on the constitutive background that makes a practice possible.

A practice therefore derives its normativity qua method insofar as it succeeds in promoting its end, but derives its normativity qua practice from its ability to constitute a shared or common will. Actions conforming the practice will therefore be an expression of a *public will*. She then distinguishes a *transgression* from a *subversion* of the practice and writes of the latter: "The subversive is one who not only breaks but usurps the law, making it into a tool of his unilateral will."¹⁰¹ In defining the abuse of practice rules as such, Schapiro succeeds in avoiding the consequentialist trap: what makes upholding participation to the practice problematic in such a case is that one ends up "serving a unilateral will", not that it makes you promote some bad end.

⁹⁸ The only reason she mentions for wanting to avoid this is to maintain the unity of moral theory, i.e. to be a deontologist and not be forced to ultimately admit that only practices have a deontological structure. Another argument can be derived from her discussion of a police administration. Although its ultimate goal may be to secure citizens, this is not what is at stake when an individual officer must choose whether to stand up to a corrupt police force: in that case, it is a conflict between the integrity of what the practice ought to be and what it is now. The dilemma, in other words, is internal to the practice.

⁹⁹ Schapiro, "Compliance, Complicity, and the Nature of Nonideal Circumstances" (2003): 342.

¹⁰⁰ Ibid: 343.

¹⁰¹ Ibid: 344.

An example illustrative of this is that of a rigged boxing match between an unbeaten, old champion, who is worth millions, and a young, rising talent struggling to catch the public eye. One might imagine a situation in which the marketing agency of the champion offers the young talent extensive marketing and an enormous amount of money for a fight in which he feints a knockout in the eleventh round. ‘Think about it,’ the marketing agency says: ‘our champion is old and has no reason smudging his career with a loss, so we could easily just reject the match, but instead we’re offering you an opportunity to enter the public eye. After this, our veteran will retire and you can claim the title in another bout. The crowd will be happy, because they get to see their champ shine one last time, and you will make millions and kickstart your career.’ From a consequentialist perspective, all this may make sense, and the interests of all parties are protected: the audience gets a fight they might otherwise not get, the champion gets to retire in style and the young talent, if he agrees, can enter the public eye and make millions. So what is harmed in the process? The answer is, of course, the sport of boxing, i.e. the practice itself, which has been rendered a sham. The practice has been made into a method to achieve fortune, fame, and spectacle. It would be simplistic to say that no harm is done if the young talent has little to no sense of honor and that all depends on his willingness to comply; rather, Schapiro would argue that the practice is no longer the expression of a public will that is constituted by practice rules, which are no more than the rules that constitute the sport of boxing.¹⁰²

Non-ideal conditions in social practices have thus been defined by Schapiro as an individual (or multiple individuals) abusing practice rules to achieve their own ends, and practices can be analyzed both in terms of their efficacy to achieve some end and their structural integrity, i.e. their capacity to express the public will, the latter occurring precisely when the practice is easily abused by individuals who wish to abuse it (such as criminals walking free too often when the burden of proof is too heavy). Schapiro distinguishes these two ways of analyzing a practice as two separate roles, namely the *social engineer* as one is concerned with creating a practice that is a method of achieving some external end, and the *reformer* as one who considers himself a participant of the practice but realizes simultaneously that the situation is

¹⁰² This example is my own: Schapiro uses the example of a sham negotiation in which the other party is using the negotiation merely to find out what you want, so that he can make an appealing offer later. Even if the outcome is favorable for all parties, the practice is still a sham. I find my example more intuitive, but perhaps it has problems that Schapiro’s example does not. Schapiro, “Compliance, Complicity, and the Nature of Nonideal Circumstances” (2003): 345.

non-ideal. It is the reformer we are primarily interested in when it comes to dilemmas of the kind described earlier.

The reformer differs from the social engineer in two respects: first, he considers himself a participant of the practice, and thus aims to live up to its constitutive rules insofar as the non-ideal conditions allow him to do so. His relationship to the practice therefore remains one of loyalty to the shared will, and not one in which the practice is a mere method to some end. Second, because it is the reformer's end to act in accord with the practice under non-ideal conditions, he may adopt the end of repairing the integrity of the practice. Thus it is not his aim to somehow optimize the practice towards its external end, but rather to restore it towards being capable of being an expression of the public will.¹⁰³

Schapiro considers it a virtue of this account that it avoids glib moralism: the agent, as a participant of a practice in non-ideal conditions, is unable to act in a way that is fully justified. The agent wishes to act in a way that expresses the shared will, but because the practice is currently not integer, no option is readily available to act in such a way.¹⁰⁴ It is not integer precisely because it is being usurped by an individual agent, rather than expressing the shared will, and it is this lack of integrity that is often associated with corruption. Considering for example the young boxer, we might find that he has multiple options that are all excusable in their own way: he may (a) accept the agreement and use his resources to restore the sport's integrity; (b) accept the agreement but violate it in the ring, so as to uphold the practice above the agreement; (c) decline the agreement altogether – and in all these options he might consider publicizing the sham as well, possibly harming the sport's or his own reputation. None of these options are ideal, because ideally the match would be fair and according to the practice rules of the sport, and this emphasizes the dilemmatic or morally ambiguous character of the situation. It is in this sense that the agent must ask “what the spirit of the law requires in this unusual circumstance,” coming up with the best available action given the situation.¹⁰⁵ Yet because the situation is non-ideal, the agent must make dirty hands, and his actions will be contaminated by

¹⁰³ Schapiro, “Compliance, Complicity, and the Nature of Nonideal Circumstances” (2003): 352-3.

¹⁰⁴ *Ibid.*, 353.

¹⁰⁵ *Ibid.*, 354.

a sense of regret that he could not simply live up to the standard set by the practice itself.¹⁰⁶

If this conception of social practices, in which a clear delineation exists between ideal and non-ideal theory, can be linked to Schapiro's conception of morality at large, then her revised Kantianism may well be equipped to respond to the case of the murderer at the door.

Met opmerkingen [PS1]: Dit is verplaatst van H3 naar H4.

4.2. The spirit of the moral law

We have seen that non-ideality in social practices, according to Schapiro, must be conceived of as an individual usurping a practice to achieve her own set end. What Schapiro fails to explain however, is how and whether this conception can be related to cases such as that of the murderer at the door. If this can be done, it may (a) settle the anything-goes complaint by providing a conception of how agents relate to the ideal morality in non-ideal circumstances, (b) provide a clear conceptual moment as to when this occurs, and (c) open the door to a revision of Kantianism that does away with problematic rigorism and marries intuition and reason, for example in the case of the murderer at the door.

Reformers and the anything-goes objection

First, how should we conceive of honesty as a social practice in Schapiro's conception? At the end of "Compliance, Complicity, and the Nature of Nonideal Circumstances," Schapiro states that she hopes to develop an account in which truth-telling is conceived of as a practice, but that more must be said to make such a comparison.¹⁰⁷ In line with her later article, this would seem to require a conception of honesty as a mode of interaction that is comparable to a social practice in that it is a constituted by certain rules. When the murderer then asks us about the whereabouts of his intended victim, he is making a move within the mode of interaction of honest communication, but he is employing that mode as a method to achieve his individual interests. Thus the situation is non-ideal and the agent is caught in a dilemmatic

¹⁰⁶ It was brought to my attention that the opposite could also be the case, that a practice such as a sport is used as a method (as conceived by Schapiro) to address a greater social issue, in comparison to which the integrity of the sport seems trivial. This is indeed quite common, such as when The Olympics are boycotted in certain countries due to human rights violations of the host. In this conception, the athlete then takes upon herself the role of the social engineer, which is lent credence due to her accomplishments as an athlete.

¹⁰⁷ Schapiro, "Compliance, Complicity, and the Nature of Nonideal Circumstances" (2003): 355.

situation in which he must decide between upholding the mode of interaction and questioning whether it is still integer.

Although I think the former conception is correct in comparing social practices (e.g. a sport or a committee) to a moral practice such as truth-telling, it does not seem to be precisely what is at stake when the murderer asks us where his victim is. It is after all not truth-telling that he is betraying when he asks this question; in fact he is being naively truthful in asking an honest question and expecting an honest answer. This seems to source the anything-goes objection, that it is unclear that we should lie to the murderer rather than attacking him physically. Yet what is at stake is the more fundamental consideration of the integrity of the moral community at large. This moral community is governed by moral principles that prohibit murder, and the murderer, in expecting to be treated with honesty, is free-riding upon those moral principles. Thus it is morality at large that he is abusing to achieve his personal ends.

In the conception of social practices, we saw that the agent who considered himself a participant of a practice that was usurped took upon himself the role of the reformer. In this role, the agent aims to both participate in the practice and restore it to the capacity of expressing the public will. It seems that this role can be transposed to the case of the murderer at the door. The agent considers himself restrained by morality, but sees that morality itself is being usurped by a unilateral will, namely the murderer. In a role comparable to that of the reformer, he responds to the situation by being (a) a participant of morality, thus guided by moral ideality, and (b) an agent in a non-ideal circumstance. In this tension between wanting to do the right thing in unfavorable circumstances, it makes sense to speak of 'living up to the spirit of the moral law.'¹⁰⁸

In response to this conception it may be argued that, once the murderer is off the moral grid, the anything-goes objection still applies, because anything may be done to restore the integrity of the moral community. Another way of conceiving of this problem is that once the murderer has betrayed the moral community, we fall back into a form of consequentialism in which our actions are mere means to the end of restoring the moral community.¹⁰⁹

¹⁰⁸ Whether the role of reformer can indeed apply to moral agents in general will be discussed in the next section.

¹⁰⁹ Schapiro, "Compliance, Complicity, and the Nature of Nonideal Circumstances" (2003): 334.

Yet this response erroneously assumes that the murderer has been removed from the moral community, which is not the case. Because morality is conceived of as an interaction rather than an action, it is possible for the murderer to betray his end of morality without us wanting to betray ours, which is precisely what makes the circumstance non-ideal. Thus the categorical imperative remains binding to agents even when others betray it, but in the latter circumstances, certain moral principles can only be lived up to in spirit, and not in letter, and therefore certain modes of interaction such as honesty become unacceptable with the murderer in favor of deeper moral considerations. The non-ideality of the circumstance forces us to deceive him, yet it does not declare him off the moral grid as Cholbi claims: for this, we would have to abandon the normative framework that underlies honesty ourselves, which was never under consideration.

An assumption that seems to be made in the anything-goes objection is that it must be deception that is proved to be especially permissible and that, if this is not proved, then murder, torture, and all kinds of horrible acts are suddenly on the table as well. This is an erroneous assumption, because from the ideal perspective there is nothing special about deception or lying in this circumstance, other than that it is presented by the empirical situation as the most intuitive and least harmful solution. It is indeed true that a realm of improvisation is entered that excuses certain acts that would otherwise be considered immoral, and indeed there can be a kind of anything-goes attitude in the sense that all options are on the table in that deliberative realm (more on this later). Yet precisely because of the double relationship of the reformer, namely as a participant and as a person in a non-ideal circumstance, can sense be made of the fact that he is still aiming to act in accord with the spirit of the moral law in this particular situation. It may however be possible that situations become so catastrophic that considerations such as killing or torturing the murderer do become conceivable. Yet in this circumstance, it would seem that in order to restore the moral community, grievous things must be done. Schapiro's account seems to make sense of this precisely because it makes the moral dilemma intelligible: we must then commit immoral actions for the sake of morality, but the gap between our actions and that ideal make the actions all the more regrettable and difficult.

4.3. Concerns and further discussion

So far, I have assumed that the role of the reformer, which is part of Schapiro's conception of social practices, can be transposed unto morality at large. I have

demonstrated that this would allow for a response to the anything-goes objection in which the moral improvisation agents in the non-ideal circumstances is linked to the moral ideal. There are, however, some questions that remain with this conception, some of which I will attempt to settle here. These difficulties arise, because morality at large is being compared to a social practice, and work must be done to distinguish where the comparison is invalid. I will address the most important concerns and discuss some other concerns for further discussion.

Reformers and protectors

It may be noted that a social practice, as an expression of a shared will, exists both as an ideal of an institution that achieves some common cause, such as police institution is meant to provide safety and criminal justice. Yet there are two ways in which such a social practice can fail to be an integer expression of the shared will. It can (1) fail ideally or conceptually, because of some lacking policy or governing principle that should exist but doesn't; or (2) it can fail empirically, not because there is anything wrong with the concept or the way it is institutionalized, but because bad people have made their way into the organization and have made it instrumental to their unilateral will. For example, if criminal organizations manage to easily get employee insiders to disperse important information to them, this might be (1) because there is insufficient policy that prevents bad people from being hired, or (2) because there needs to be a purging of bad people from the organization (or both).

In the role of the reformer, it seems both (1) and (2) are eligible options when wanting to restore some police organization. It seems that the first option is most appropriately called reform, because there is a rethinking of what the form of the institution should be, disregarding its empirical circumstance. The second option is more properly called a *protection* of the institution, which is in principle integer but must be protected from bad individuals who are abusing it and subjecting it to their will.

In the case of morality at large being usurped, which is what is at stake in the case of the murderer at the door, the first option, that of reform, is never available, because there cannot be some lacking piece of policy from morality itself. Morality is, by definition, the ideal concept of the laws that constitute the integrity of the actions and interactions of and between autonomous agents – it cannot therefore itself be corrupt or in need of reform. It is more appropriate, especially when the moral community is considered as the collection of autonomous agents, to speak of the role

of *protector* that is assumed when morality is being usurped as it is in the case of the murderer at the door.

The status and spirit of honesty in Kantianism

Another concern is the status of truth-telling in Kantianism. The discussion into rigorism and the necessity of fixing it occurs because Kant's moral theory lends itself for conceptions in which we approach the moral status of e.g. deception with emphasis on the consistency of the action. Honesty then receives special status, because one can never consistently tell a lie for what Cholbi calls a double standard: we depend on the truth for the lie to be believed.¹¹⁰

Yet intuitively, a lie to prevent a greater harm (a mitigating lie) is not so problematic. Particularly in the circumstance of the murderer at the door, it seems that honesty is completely off the table and we might consider not lying to the murderer to be the less defensible course of action. This prompts Cholbi to want to argue for a duty to protect the murderer's victim, and prompts Schapiro to ask what it is that is right about honesty (rather than emphasizing what is wrong about deception).

Yet by conceiving of morality as something that governs interactions rather than actions, this special status of honesty, which before it received from concerns related to the consistency of the action, can be made sense of in revised Kantianism. Honesty becomes a matter not of mere truth-telling, but of considering what information we owe to one another given that we respect the autonomy of other agents. The information we owe to one another seems to be the spirit of honesty proper, and if we were to act in the spirit of honesty when a murderer inquires about the location of his victim, we owe him no information in that circumstance because he has betrayed deeper moral considerations.

Although this may not settle the matter of honesty in revised Kantianism entirely, it may open the discussion towards a more nuanced view of honesty in deontology, which can be had once the matter of rigorism has been settled.

Topics for further discussion

There are additional concerns to this conception, that either have either ran throughout this discussion or that arise in this revised conception of Kantianism.

¹¹⁰ Cholbi, "The Murderer at the Door: What Kant Should Have Said" (2009): 44.

One such concern is that, although a relationship between ideal morality and non-ideal improvisation has been proved, it has not been made entirely clear what that relationship, or proper moral improvisation in non-ideal circumstances, would look like. Schapiro mentions that agents “are forced to improvise”, yet what kinds of principles apply in this realm (if any) is itself subject for discussion.¹¹¹ In this sense, the fact that we aren’t disconnected or removed from the moral community in non-ideal circumstances is of little help: there is still no real framework that excuses for example deception over physical restraint, or answering murder with murder. This concern makes apparent why we are speaking of an excuse and not of a justification. If the murderer is, for example, invited in and consequently killed by the agent, we might consider that less excusable than the lie, but nonetheless there is a degree of sympathy for responding that way (perhaps depending on the circumstances). On the other hand, there is no justification for a killing that follows straight forwardly from the categorical imperative: it is only in such exceptional and non-ideal circumstances that we would excuse this behavior.

Relatedly, within this realm of improvisation there seem to exist considerations of proportionality that are not explicit in this account. As said, lying is a much more excusable response than murder, but there are no clear indicators as to why this is the case and whether this may provide normativity for our conduct when forced to improvise. Additionally, there are more empirical parameters, such as the duration of the improvisation or trusting the murderer after the situation has been dealt with – much of these matters are simply left to our best judgment.

Additionally, although I have discussed some of the matters regarding how morality at large can be compared to a social practice, it is not entirely clear how much this answers. In this thesis, the particular circumstance that has been discussed is one in which a murderer usurps morality to forward his own ends, but this does not answer all dilemmatic situations. One such situation is described by Schapiro in a footnote, which discusses the example of a situation in which an agent might lie to Jack about Jill’s location, because Jack is a blabbermouth and Jill is potentially being stalked.¹¹² In this circumstance, one can imagine the non-ideality of the situation extending towards the factual circumstance that Jack is a blabbermouth, yet it remains a form of deception that deontology has difficulty explaining. A possible solution is to

¹¹¹ Schapiro, “Kantian Rigorism and Mitigating Circumstances” (2006): 47.

¹¹² Schapiro mentions this as a topic for discussion in footnote 29 at the end of her article. Ibid 56.

introduce more factors that could render situations non-ideal, such as the matter of how well someone handles sensitive information: it is perfectly intuitive to state that one would have liked to trust Jack with the information of Jill's location, but that circumstances (e.g. Jack being a blabbermouth and, possibly, not knowing that Jill was being stalked) forced one to lie. Yet to make sense of such considerations, more work must be done to make sense of this realm of non-ideal circumstance and its appropriate forms of deliberation.

Finally, more work must be done to flesh out the conception of deontology as the morality of interaction, rather than of action, and what the consequences are for the formulations of the categorical imperative to conceive of morality as a shared standpoint. The fruit of this already lies within Kant's moral theory, particularly the kingdom of ends formulation of the categorical imperative, yet requires a deviation from the rigorist elements of his philosophy, which overstates the relevance of moral consistency. There is, of course, space for a right kind of rigorism, which is pivotal to deontology, namely rigorism of the orientation of the will. The latter form of rigorism is one which ensures that agents live up to the spirit of morality even in non-ideal circumstances, when ideal theory is not fully informative and agents are forced to improvise.

Conclusion

The aim of this thesis has been to consider the validity of Kantian rigorism in non-ideal situations such as the case of the murderer at the door, and to assess whether Schapiro's revised Kantianism can provide a satisfying response. In chapter 1, I have argued that it is plausible to interpret Kant as a rigorist. Kant was at least a rigorist in the sense that he believed that there are only two orientations of the will, namely a good and a bad one, the former with respect for the moral law and the latter out of inclination. Yet when this rigorism is combined with his stringent view on the case of the murderer at the door in *On the Supposed Right to Lie*, an interpretation arises in which agents are instructed to abide by the ideal set by the formulations of the categorical imperative without regard for mitigating circumstances. The attempts of Kantian apologists to offer a historical explanation for that essay did not make the problem go away: it is not clear when agents are allowed to deviate from the ideal prescriptions of the categorical imperative.

Cholbi's approach, discussed in chapter 2, was to build upon Kant's moral theory, but to deviate where necessary to provide a satisfactory response to the case. His aim was to argue for a duty to save the victim, rather than merely a permission to deceive the murderer. He created space for the permissible deception by arguing that there exist permissible lies when the immoral actions of the murderer are incorporated into the universal law formula, and that we are permitted to thwart immoral actions in the kingdom of ends. Through the deontic symmetry thesis, he argued that the duty of self-preservation applies similarly to others, and that we therefore have an obligation to save the victim. The problem with this approach however was that, while Cholbi's ambition was to justify deception without appealing to non-ideal theory, by incorporating the immoral intentions of other agents into our maxims he did precisely that. As a result, his position became open to a myriad of objections, one of which ironically being that anything can be universalized once we allow the murderer's intentions into our maxim when testing it against the universal law formula.

Schapiro therefore suggested to revise Kantianism. In summary, her approach was reconceive of honesty as a mode of interaction, and to conceive of morality as a fundamental practice constituted between agents. This moral community is governed by the categorical imperative, and it is within it that moral action is possible, i.e. that individuals become agents, similar to how in Kant's moral theory an agent can only be free when she acts in accordance with the moral law. In her account of practices, it became apparent that it is possible for an agent to usurp a practice by abusing it for personal interests. In the case of the murderer at the door, this account can be extended to the moral community at large, which the murderer abuses when he requests honesty to further his personal, immoral agenda. This is a clear moment at which agents enter into non-ideal situations, namely when the moral community is abused to further the personal ends of another agent. Agents then do not abandon the ideals of the moral community, but in the face of mitigating circumstances they are forced to improvise and live up to its ideals to their best ability.

Schapiro's revised Kantianism does not entail an abandonment of rigorism altogether: there is a right kind of rigorism which seems quite fundamental to deontology, and a wrong kind of rigorism. The latter has been the primary topic of discussion, because without space for non-ideal theory it seems to follow from Kant's deontic rigorism. It is the kind of stringency and lack of regard for mitigating circumstances that has been discussed throughout this paper. Yet the right kind of rigorism, rigorism of the will, is the one which states that there are only one good

orientation of the will, namely with respect for the moral law, and that all else is natural inclination or self-love. Schapiro's revised Kantianism does not need to abandon this form of rigorism; in fact, the good orientation of the will seems pivotal in acting in the spirit of the categorical imperative, because our improvisation must not be influenced by natural inclination when we are forced to deviate from the ideal of the moral law.

References

- Benton, Robert J. "Political Expediency and Lying: Kant vs Benjamin Constant." *Journal of the History of Ideas*, Vol 43. No. 1 (Jan. – Mar., 1982): 135-144.
- Cholbi, Michael. "The Constitutive Approach to Kantian Rigorism." In *Ethical Theory and Moral Practice*, Vol. 16, No. 3 (2013): 439-48.
- Cholbi, Michael. "The Murderer at the Door: What Kant Should Have Said." In *Philosophy and Phenomenological Research*, Vol. 79, No. 1 (2009): 17-46.
- Johnson, Robert & Cureton, Adam. "Kant's Moral Philosophy." In *The Stanford Encyclopedia of Philosophy* (Fall 2017 Edition). Edited by Edward N. Zalta. URL: <https://plato.stanford.edu/archives/fall2017/entries/kant-moral/>
- Kant, Immanuel. *Groundwork of the Metaphysics of Morals: A German-English Edition*. Edited and translated by Mary Gregor & Jens Timmermann. New York: Cambridge University Press, 2011.
- Kant, Immanuel. "On the Supposed Right to Lie from Altruistic Motives." In *Critique of Practical Reason and Other Writings in Moral Philosophy*. Edited and translated by Lewis White Beck. Chicago: University of Chicago Press, 1949, 346-50.
For the German text: Kant, AA VIII: Abhandlungen nach 1781, url: <https://korpora.zim.uni-duisburg-essen.de/kant/aa08/Inhalt8.html>
- Kant, Immanuel. *Religion Within the Boundaries of Mere Reason And Other Writings*. Edited and translation by Allen Wood & George Di Giovanni. Cambridge University Press, 1998.
- Mahon, James E. "Kant on lies, candour, and reticence." *Kantian Review* 7 (1) (2003): 102-133.
- Paton, Herbert James. "An Alleged Right to Lie." *Kant-Studien* 45, 1-4 (1954): 190-203.
- Rawls, John. *A Theory of Justice*. Cambridge, Massachusetts: Belknap Press of Harvard University Press, 1971.
- Scanlon, Thomas M. *What We Owe to Each Other*. Cambridge, Massachusetts, and London, England: The Belknap Press of Harvard University Press, 1998.

Schapiro, Tamar. "Compliance, Complicity, and the Nature of Nonideal Circumstances." *The Journal of Philosophy*, Vol. 100, No. 7 (2003): 239-55.

Schapiro, Tamar. "Rigorism and Mitigating Circumstances." *Ethics* 117 (2006): 32-57.

Valentini, Laura. "Ideal vs. Non-Ideal Theory: A Conceptual Map." *Philosophy Compass*, 7 (9) (2012): 654-64.

Williams, Bernard. *Problems of the Self. Philosophical Papers 1956-1972*. Cambridge: Cambridge University Press, 1973.