

Claremont Colleges

Scholarship @ Claremont

CMC Senior Theses

CMC Student Scholarship

2023

A Defense of Kantian Ethics Against Rigorism

Leyna Hong

Follow this and additional works at: https://scholarship.claremont.edu/cmc_theses



Part of the [Applied Ethics Commons](#), and the [Ethics and Political Philosophy Commons](#)

Recommended Citation

Hong, Leyna, "A Defense of Kantian Ethics Against Rigorism" (2023). *CMC Senior Theses*. 3187.
https://scholarship.claremont.edu/cmc_theses/3187

This Open Access Senior Thesis is brought to you by Scholarship@Claremont. It has been accepted for inclusion in this collection by an authorized administrator. For more information, please contact scholarship@cuc.claremont.edu.

Claremont McKenna College

A Defense of Kantian Ethics Against Rigorism

Submitted to
Professor Andrew Schroeder

By
Leyna Hong

for
Senior Thesis
Fall 2022
December 5, 2022

Table of Contents

Introduction.....	1
Chapter 1: Kantian Ethics and the Commitment to Truth	6
Categorical Imperative	7
Formula of Universal Law	8
Formula of Humanity	10
Kingdom of Ends	12
Commitment to the Truth in Kant’s Theory.....	13
Response to the “Murderer at the Door” Hypothetical	15
Chapter 2: Korsgaard’s Ideal vs Non-Ideal Theory	18
Chapter 3: Schapiro’s Constitutive Approach.....	24
Chapter 4: Cholbi’s “Anything Goes” Objection to Schapiro.....	34
Chapter 5: Defense of Schapiro	38
Further Considerations	42
Bibliography	45

Introduction

Among the most popular ethical theories, Kantian ethics has proven to have its advantages. Like other deontological theories (ones based on rules or principles), Kant's ethical theory shares the advantage that the moral principles that constitute the theory provide clear rules as to how one should act morally. There are clear do's and don'ts in deontological theories that delineate what is morally required and what is morally prohibited. Kantian ethics stands out from other deontological theories because its principles are derived from the respect of other people as rational beings. In other words, people are at the heart of Kantian ethics, so the moral rules put people, including ourselves, as the good with the most moral value. Other ethical theories may place intrinsic value on our specific desires, such as happiness, but ultimately, the reason why we choose to value or pursue our desires is because we value ourselves. In consideration of these points, Kantian ethics seems to offer an ethical theory that can clearly guide our actions while appealing to our values.

While Kantian ethics provides clear, intuitive rules that serve well to guide our everyday actions, there is criticism and debate on how Kantian ethics addresses moral dilemmas. In a moral dilemma, there does not seem to be a clear course of action because with any choice we make, we transgress one of Kant's moral principles. Let us consider one moral principle that Kant develops — telling the truth. Normally, we would like to say that “honesty is the best policy,” but we recognize that there may be some moral dilemmas where our intuition says otherwise. We likely find reasons to lie in our daily lives, especially when we tell white lies to avoid hurting someone's feelings. For instance, let's say that you go to a coffee shop with your friend. As you go to pick up the

drinks while your friend waits for you at a table, the barista tells you that they think your friend is unattractive and makes unnecessarily rude comments about your friend's appearance. You call the barista out for the comments, but when you return to your table, your friend comments on how attractive the barista is and asks if you think they should try flirting with the barista. Obviously, you know the attraction is not reciprocated, but sharing with your friend the barista's comments may cause unnecessary harm. Your friend is unlikely to ever see the barista again, and sharing the barista's comments would only hurt their feelings. Rather than tell your friend how the barista feels, you lie about the situation to prevent your friend from flirting. Perhaps you lie and say that they smelled like cigarettes because you know your friend does not like smokers, or lie about how you have somewhere else to be and they don't have time to flirt with the barista. We all likely have similar situations where we tell small white lies to spare a person's feelings or avoid unnecessary conflict because telling the truth in these situations may seem marginally beneficial, if not destructive.

In the barista example, some people may argue that our friend deserves to know the truth regarding what someone said about them. We do not have the right to decide what they should or should not know, and our friend can learn to not let other people's comments affect their self-esteem. We could also mitigate the negative effects by being supportive and reminding them that the barista's comments are not true. Lying, in this case, is a way for us to make the situation easier to deal with, but that does not mean it's the right action.¹ Let us consider, then, another example where the moral dilemma is more

¹ Here, I am not arguing for whether or not we could morally justify white lies. I am simply showing that while we may see the benefit to lying, as understood intuitively with common white lies, the argument could be made that the lies in these situations are not necessarily morally right.

extreme. Let's say that your friend is running away from a murderer and has sought refuge in your home. The murderer shows up to your door and asks if you know where your friend is. Most people would say that we would lie in the situation to protect our friend.

Our intuition tends to be that lying is not as bad as killing. In our day to day, we find more ways to justify lying (to spare someone's feelings, mitigate conflict, etc.), but rarely can we find ways to justify killing (mainly just self-defense). We would also likely not be as committed to saying someone is a bad person if we knew they lied before as opposed to if they killed someone before. As a result, it may seem like we have a *ranking* of some kind for moral wrongdoings — while lying and murder are both at default considered wrongdoing, lying is not as bad as murder. In the “murderer at the door” scenario, it seems as though our only two options are to lie or be complicit in a murder. The appropriate response seems simple and obvious — do not tell the murderer the truth because murder is worse than lying. However, this opens up a few questions for Kantian ethics. First, how could it be that one moral principle is more significant or fundamental than another (i.e., killing feels more fundamentally wrong than lying)? Also, to say that we should lie to the murderer is to say that the right action in this case is to not tell the truth. If this is the case, then it opens the questions of whether Kantian ethics can make *exceptions* to the moral principle of telling the truth, and if so, how will these exceptions fit into the rest of the theory.

Kant actually shared his response to this “murderer at the door” scenario in his letter “On the Supposed Right to Lie” (which we will look at more closely later), and he

However, the next case I introduce, the “murderer at the door” case, presents a case where lying seems to be morally justifiable.

controversially concluded that the agent should tell the murderer the truth (1797). It is important to note that this response was controversial even in the time he wrote it. His response was to a letter written by Benjamin Constant who posed this “murderer at the door” scenario because he believed it to be a case where the right action would make it clear that exceptions are needed for the moral principles (Constant 1797). Kant stubbornly stuck to his theory and refused to give the response that Constant wanted. Since Kant published this response, the challenge of *rigorism*, or the seeming inflexibility of Kantian ethics to moral dilemmas, has become the focus of philosophers seeking to dismiss Kantian ethics and those hoping to defend him.

Given the strong appeal of the rest of Kant’s ethical theory, I hesitate to dismiss Kant’s ethical theory altogether because of the rigorism. I would like to first seek a charitable interpretation of Kant before looking elsewhere for another ethical theory, so my goal in this paper will be to pursue the best defense I can for Kantian ethics against rigorism. In the first chapter, I will discuss Kant’s ethical theory and the letter in which he provides his controversial response. In the next two chapters, I will present two charitable interpretations of Kant that try to reimagine Kantian ethics to permit exceptions. The first approach that I will discuss will be Christine Korsgaard’s, where she proposes to distinguish between ideal vs non-ideal circumstances as a way to permit exceptions. The following chapter will be on Tamar Schapiro’s constitutive approach where she believes the solution to rigorism can be found in interpreting honesty as constituting multiple facets beyond simply telling the truth. Then, I’ll discuss Michael Cholbi’s challenge that Schapiro’s approach is too permissive. Finally, I will provide a defense of Schapiro’s

approach against Cholbi's objections, as well as some questions that can be further considered regarding the moral differences behind lying and killing.

Chapter 1

Kantian Ethics and the Commitment to Truth

Before we can discuss the challenges of rigorism, we must have a foundational understanding of Kant's ethical theory. Accordingly, in this chapter, I will discuss the foundations that Kant lays out in his *Groundwork for the Metaphysics of Morals* (1785). While the *Groundwork* does not cover all of his metaphysical theory on morals, it develops the fundamentals of his theory. There are other works that expand on his moral theory, such as *The Metaphysics of Morals* (1797), but the *Groundwork* holistically and sufficiently discusses the elements of his theory we need to understand the challenge of rigorism.

Kant begins by considering different examples or hypothetical situations to understand how we decide which actions have moral worth. He considers how we intuitively understand moral worth and seeks to find the overarching idea or principle that unites them. He draws a distinction in types of *maxims*, or motivating reasons behind their actions. Kant argues that actions lack moral worth when the agent only chooses to do something out of inclination, or desire, because this is a self-serving purpose. Instead, acts only have moral value when they are done from a sense of duty; the agent does the act because she thinks it's required of her to do so (1785 4:398-4:400). For example, both agents A and B may volunteer for a homeless shelter. Agent A does it because she enjoys making others happy. Agent B does it because she thinks it is required to help others. Kant would say only agent B's actions have moral value because agent B acts out of duty, whereas agent A's actions are motivated by inclinations, or desires. Kant argues it is

essentially the principle upon which she acts that determines the moral value of an act, and the sense of duty comes from “respect for the law” (1785 4:400). In other words, to act from a sense of duty is to comply with a law rather than act merely on our desires. This empirical examination serves as the intuitive basis for why our theory should be driven by laws, or moral principles. In consideration of what kind of law moral laws should be, Kant recognizes that these laws should be applicable in all situations. As a result, he concludes that we should act in a way where “I could also will that my maxim should become a universal law” (1785 4:402). To sum up, our main takeaways from the first chapter of *Groundwork* are that from examining empirical examples, our intuitions indicate to us that the moral worth of our actions comes from 1) examining the maxims that motivate our actions, 2) actions motivated by a sense of duty as opposed to inclinations, and 3) choosing maxims that can be universalized.

Categorical Imperative

The first chapter provides the intuitions to help us understand moral laws, but in the rest of *Groundwork*, Kant develops the theory behind his argument. From hereon, he develops his arguments for the thesis of his theory in which he calls the *categorical imperative*, which states that an agent should “act only according to that maxim through which you can at the same time will that it become a universal law” (1785 4:421). The Categorical Imperative (CI) can be understood in three ways, or as having three forms — the Formula of Universal Law, Formula of Humanity, and Kingdom of Ends. These three forms serve as three different ways of arguing for the same ultimate CI. We have three

different arguments because they are all essential to the CI — they all highlight a unique and essential quality of the CI and work hand-in-hand to develop the CI as a whole.

Formula of Universal Law

In trying to understand the nature of moral rules, Kant questions what dictates what is demanded of us morally. In other words, how do we come to develop and understand what morality demands of us? To answer this question, Kant first considers two types of demands, and from hereon, we can understand a demand as being an *imperative* (1785 4:414-4:420). First, there are hypothetical imperatives that follow the structure of “If I will X, then I ought to will Y.” For example, if I will a healthy life, then I ought to exercise. The hypothetical imperative is conditional because we do not value our action (from the example, exercising) for the sake of itself, but rather to accomplish something else (e.g., health). Moral principles cannot be hypothetical, however, because they must be unconditional. Moral principles cannot call upon some higher law or order to guide it because the moral principles are what guide us. Their value must be intrinsic; therefore, moral principles must be laws themselves. Because laws are meant to apply to everyone and be universal, we can understand these maxims as being universal as well. As a result, we come to understand moral principles as a *categorical*, as opposed to a hypothetical, imperative to illustrate their unconditionality. Kant proposes the *Formula of Universal Law* as a test to see whether or not an act meets the categorical imperative, and thus, is morally permissible.

The Formula of Universal Law dictates that the universal imperative of duty can be expressed as “so act as if the maxim of your action were to become by your will a

universal law of nature” (1785 4:421-4:424). To use the Formula of Universal Law, take the maxim that motivates the act, universalize it, and see if it forms a contradiction. There is an unresolved debate in what type of contradiction (logical or practical) that Kant means, but for the purposes of this paper, I will use the understanding of “practical” contradiction — “the universalized maxim contradicts itself when the efficacy of the action as a method of achieving its purpose would be undermined by its universal practice” (Korsgaard 2004 135). In considering whether universalizing the maxim forms a contradiction, we are asking ourselves here, “In a world where everyone acted in the same way as I did with the same maxim, would I still be able to accomplish the desired goal of my action?” If it does not make sense or is impossible, this would mean that your maxim forms a contradiction when universalized.

Universalizing a maxim may sound strange or non-intuitive, so let’s consider a moral duty that Kant develops through the Formula of Universal Law — the moral duty of keeping one’s promise. Let’s say you borrow money from someone and promise to pay them back. You make this promise, however, with no intention of actually keeping it. In this isolated case, we are able to accomplish our goal so long as the person whom we borrow money from does not know of our true intentions. The other person believes our false promise that they are loaning us money for us to eventually return it, so we are successful in getting their money despite our plans to not return it. We would not be successful, however, if our maxim of not keeping our promises to return other people’s money was universalized. If we were to universalize the maxim “I will break the promise of borrowing money to steal other people’s money,” that would mean that when anyone makes a promise to return someone’s money after borrowing it, the promise will always

be broken. Accordingly, no one would ever agree to such a promise because they know that they will not get their money back. As a result, having the maxim universalized would make the goal of the maxim impossible. This impossibility is where we see a practical contradiction — the universal practice of our maxim would make it so no one would be able to accomplish the goal set out by that maxim.

The principle of telling the truth can be developed on a similar line of thinking to keeping a promise. Let's take the maxim, "I will lie and deceive others to achieve my ends." In universalizing the maxim, it would be the case that everyone deceives others to achieve their personal ends. If these ends rely on other people, then you would not be able to achieve your ends because no one would be fooled by your deception. People will not trust anyone else and will not fall for the lies that you tell to achieve your ends. As a result, lying violates the Formula of Universal Law due to the practical contradiction universalizing it would create, so we can see how lying is morally prohibited under the Formula of Universal Law.

Formula of Humanity

The next form of the CI Kant develops is the Formula of Humanity. First, Kant takes a step back to see how we form maxims to begin with. He argues that we do so by considering an end (1785 4:427). The ends that we typically use to explain our actions have relative, not absolute, value. For example, we value learning because we desire knowledge, or we value art because we enjoy how it appeals to our senses. The desires for knowledge and art are relative to our nature, so Kant considers these as *relative* ends (1785 4:427). These relative ends can only serve as grounds for hypothetical imperatives

because the ends are pursued for a higher purpose — namely, our desire of the end is what gives them their worth. This is problematic for Kant because as we had seen in the development of the categorical imperative, our moral principles should be grounded in something that is absolute and unconditional. Here, Kant seeks to find something that is *an end in itself* (1785 4:428). In considering how relative ends only acquire their worth given their importance to us rational beings, Kant recognizes that we do what is important to us because we see ourselves as important. We see *ourselves* as an end. We don't see ourselves as having conditional importance relative to a higher purpose, so Kant concludes “a rational nature exists as an end in itself” (1785 4:429).

We should not only see ourselves as an end in itself, but we should also treat anyone with a rational nature as an end in themselves. To not treat a rational being as an end would be to treat them as *a means to an end* instead. That is, we see the other person as a tool to be used to achieve another end. However, because rational beings ultimately have absolute worth, we should never treat another rational being as a means. The Formula of Humanity is as follows — “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” (1785 4:429). To act according to the Formula of Humanity, Kant implores us to develop our rational capacities and promote each other's chosen ends by allowing them to make their own decisions (1785 4:430-4:432). Let's consider the duty developed earlier in which Kant says we must not make false promises but now under the scope of the Formula of Humanity. The end in which the liar is trying to accomplish is the possession of someone else's money. The liar recognizes, though, that the other person will only give them the money if they believe it will eventually be returned. To accomplish their end,

the liar gives false information about their intentions to return the money. By giving false information, the liar manipulates the other person's rational capabilities to achieve their end. As a result, they see the other rational being as a mere means to an end.

Kingdom of Ends

To answer the question of how we are motivated to act morally, Kant develops his argument for the *Kingdom of Ends*, the last form of the CI. Because we are the ones who make the maxims that motivate our actions, we are the ones making the laws that get passed through the Formula of Universal Law. The only laws that we can make, however, must align with what humanity values given that we must promote each other's chosen ends (1785 4:430). As a result, as rational beings, we are the ones making and legislating the law, and we can understand "the will of every rational being as a universally legislating will" (1785 4:431). Unlike some laws that are legislated externally and we are motivated to follow it from fear of repercussions, we legislate it ourselves. Legislating these laws ourselves means we endorse the law and believe we ought to act that way. By seeing ourselves as legislating laws, we can envision ourselves as legislators in what he calls the Kingdom of Ends. This is a theoretical, ideal society where the laws that we make ourselves respect everyone as ends in themselves, and everyone is a rational being that can choose their own ends. To understand the Kingdom of Ends in relation to the two other forms of the CI, the Kingdom of Ends is a place in which your maxim *does* become universal law, and the laws reflect the respect of other rational beings from the Formula of Humanity. The Kingdom of Ends is where everyone treats one another with respect, and the maxim that motivates your actions should be endorsed and adopted universally.

Commitment to the Truth in Kant's Theory

Now that I've laid out the basics of Kant's theory, let's consider what Kant's views may be on the relationship between the moral principle of honesty and the moral prohibition against lying. On the one hand, lying violates the CI, and it seems as though there is no room to make exceptions for lying. As aforementioned, lying can be understood as violating the Formula of Universal Law and Formula of Humanity. It creates a practical contradiction when trying to universalize the maxim, and perhaps more importantly, lying treats others as mere means to an end, and thus, disrespects their status as a rational being. Furthermore, there seems to be a lot at stake for the moral principles, so exceptions for lying seem impossible for Kantian ethics. Kant continuously emphasizes the importance of the unconditionality and intrinsic value of the moral principles. The wrongness or rightness of the moral principles seem like they must be inherent, so the question arises if these principles can ever be considered less wrong or less right given the particular circumstances. In other words, if the moral principles are unconditional, can the circumstances ever change the moral value of an action? It seems that if Kant wants to preserve the importance of the unconditionality of moral principles, then circumstances cannot ever be considered as a condition for determining whether exceptions can be made to lying as a moral prohibition.

On the other hand, one may point to specific points of Kant's writings where he seems to endorse the idea of exceptions. In the fourteenth footnote of the *Groundwork*, Kant understands perfect duty as "one that allows of no exception to the advantage of inclination." This clause may be taken as evidence that Kant does not dismiss exceptions

altogether but rather only those exceptions that are driven by desire as opposed to duty. One of the perfect duties he develops is that of keeping one's promises. Telling false promises, as discussed earlier, can be likened to that of lying, so we can conclude that Kant would also see telling the truth as a perfect duty. We may take this as a possibility that there can be exceptions made to the duty of telling the truth, but Kant does not explain further how we can determine when exceptions can be made to the moral truths.

It is also important to note a line from Kant's *Lectures on Ethics* in which it can be argued that Kant thinks lying is permissible for self-defense reasons. He writes:

“... if we were to be at all times punctiliously truthful we might often become victims of the wickedness of others who were ready to abuse our truthfulness. If all men were well-intentioned it would not only be a duty not to lie, but no one would do so because there would be no point in it. But as men are malicious, it cannot be denied that to be punctiliously truthful is often dangerous... if I cannot save myself by maintaining silence, then my lie is a weapon of defense.” (1775-1780 228)

Kant recognizes that we do live in a world where people do not see others as ends in themselves and can commit wrongdoing against us. We shouldn't allow ourselves to be hurt by others who try to take advantage of our truthfulness. As a result, Kant does in fact allow lying to some degree, but it is still unclear if lying is permissible under other circumstances. We could also start to question, however, the integrity of Kant's theory. If the moral principles are unconditional, how could it be that there are some conditions in

which the circumstances, such as the need for self-defense, allow us to deviate from the moral principles? There is a balance that needs to be found in defending Kant against rigorism while still preserving the integrity of his theory, and we will continue to ask this question throughout the paper.

Response to the “Murderer at the Door” Hypothetical

To consider what Kant has to say about exceptions for lying, let’s now revisit the “murderer at the door” case posed earlier in the introduction. Benjamin Constant challenges Kant’s ethical theory in his article “On Political Reactions” by proposing a situation he believes should be an exception to the principle that one must tell the truth (1797). In this situation, your friend is running away from a murderer and has sought refuge in your home. The murderer shows up to your door, but your friend is hiding in your home. When he asks you if you know where your friend is, do you tell him the truth or lie? Constant proposed this situation because he believed the intuitive response is that you would lie to protect your friend. Constant may have used a friend as the subject of the example to strengthen the intuition, but we can even consider a stranger or relate this to a real example back to World War II. Many Jews sought refuge in other people’s homes, and Nazis would in fact go to people’s homes searching for Jews in hiding. With this example, it seems obvious that saving the lives of the Jews have more moral value than always telling the truth. The harm you may cause the Nazi by not telling the truth is incomparable to the harm you would cause to the Jews by exposing them and ultimately sentencing them to death. To argue that there is no exception to the principle of always telling the truth would thus seem absurd. Constant argues that it would make any society

“impossible if that principle [telling the truth] was taken singly and unconditionally” (1797 123).

In his response “On the Supposed Right to Lie Because of Philanthropic Concerns,” Kant argues that Constant’s “murderer at the door” scenario does not demonstrate a need for an exception to the principle. Kant states that, to be truthful is “a sacred and unconditionally commanding law of reason that admits of no expediency whatsoever” (1797 2). First, he reiterates that telling the truth is not just a duty, but a duty to all of humanity. He acknowledges that in lying, you do “no wrong to him who unjustly compels me to make a statement,” but by lying you “do wrong to duty in general” (1797 1). Kant goes further to say that there is a difference between doing harm and doing wrong to another. He accepts that harm is done to the friend, but this harm is accidental in the sense that the one who tells the truth is not responsible. This is because the murderer is the one who ultimately causes the harm, and the agent has no control on the outcome. The agent is also “not at all free to choose in such a case” because it is an “unconditional duty which holds in all circumstances” to tell the truth (1797 2).

His response can seem surprising given that we have some evidence of him seemingly claiming otherwise — that duties cannot have exceptions regarding inclinations, and in particular, lying is permissible in the face of self-defense. At the same time, though, his response seems to defend his theory against our questions of the unconditionality and intrinsic values of our duties. However, it may feel as though Kant is defending his theory too rigidly. Our intuitions, as discussed earlier, tell us that we *should* lie in order to protect an innocent person from a senseless murder. Kant’s response may seem to be too focused on defending his theory that he has lost sight of the intuitions

that tell us that telling the truth would somehow feel *wrong* in this case. For the next few chapters, I'll explore the responses from philosophers who try to provide explanations as to how Kant's theory could help us arrive at the conclusion that it is morally permissible for us to lie in this case.

Chapter 2

Korsgaard's Ideal vs Non-Ideal Theory

I'll first consider Christine Korsgaard's attempt to reconcile Kant's "On the Supposed Right to Lie" response and our intuitions that lying to the murderer should be permissible. In her paper "The Right to Lie: Kant on Dealing with Evil," she first acknowledges how Kant may seem to contradict himself in his response regarding lying (Korsgaard 1986). In trying to understand what the CI would dictate the right course of action be in the "murderer at the door" scenario, one may be able to use the Formula of Universal Law to come to a conclusion that lying to the murderer can be permissible. As aforementioned, lying was one action that Kant himself demonstrated cannot be permissible because it does not meet the Formula of Universal Law test. The goal of lying is to deceive, but if everyone lied, no one would be deceived because they would know not to trust anyone. Korsgaard recognizes how this situation is different from simply just lying. If the murderer is asking you where your friend is, Korsgaard takes this as a sign that the murderer doesn't think you know who he actually is or his intentions.² Now let's universalize the maxim that one can lie when asked by a murderer the whereabouts of their victim who is innocent. Unlike universalizing the maxim of lying in a typical case, Korsgaard argues that the murderer would not suspect they are being deceived because the murderer believes you are unaware that he is in fact a murderer. In other words, even though everyone is aware of and follows the maxim that you are

² An argument can be made that his asking of your friend's whereabouts does not necessarily indicate he believes he is deceiving you. Perhaps the murderer knows you are a Kantian and will not lie to him regardless of his status as a murderer. Korsgaard thinks it is more likely than not that the murderer is hoping to deceive us, but she gives an explanation of what we should do if the murderer is not deceiving us. This explanation will be discussed later in the paper.

permitted to lie to murderers, by asking you about his victim's whereabouts, he has shown that he believes you are unaware of his intentions, and thus, is not worried about the universalized maxim.

Under the Formula of Universal Law, Korsgaard believes that the argument could be made that lying can be permissible if there is already deception in play by the other party. The maxim that one can lie to the murderer can be universalized because the murderer hopes that others are unaware of his status as a murderer. However, Korsgaard argues that while the Formula of Universal Law can be used to show a contradiction in Kant's response, it would be a mistake to look at the Formula of Universal Law in isolation when there are two other parts of his theory — the Formula of Humanity and Kingdom of Ends. By examining these two parts of his theory, we can understand why Kant would make the argument for telling the truth.

Korsgaard mainly focuses on the Formula of Humanity in explaining Kant's rationale. In particular, she focuses on this specific statement he makes from the *Groundwork of the Metaphysics of Morals* regarding why we cannot use others as a mere means:

“For he whom I want to use for my own purposes by means of such a promise cannot possibly assent to my mode of acting against him and cannot contain the end of this action in himself” (1785 4:430)

Let's first look at what Korsgaard takes as “cannot possibly assent to my mode of acting against him” to mean. First, it's important to note that *cannot* is different from *would not* or *should not* because rather than simply not desiring to assent to the action, it

is *impossible to*. Consider a false promise someone makes about paying me back after borrowing some money from me. If I were to express assent, I'm assenting to the promise that they will borrow but eventually return my money. However, there is no such promise because the other person does not plan to return my money, so the assent I give is for a promise that does not actually exist. I'm not truly assenting to the false promise. Even if I were aware that their promise is false *and* I wanted them to keep the money, I am not assenting to the false promise that they will pay me back. Instead, I am doing something else entirely, such as gifting them the money. As a result, it is impossible to assent to acts where what is expressed is different from what is actually happening. In cases of deception, we are unable to give our assent to the deception.

If we cannot assent, this means we are not free to choose how to respond. We are not given a say in the outcome of events; instead, we are used as a tool to achieve a certain outcome. In Kantian terms, if a person cannot dictate the end, this means they are being treated as a mere means to an end. This leads to the last clause of the quote — “[he] cannot contain the end of this action in himself.” By being given a false promise, we are not free to choose to contribute to the realization of an end. In the case of borrowing money, I believe I am contributing to the other person's borrowing of my money, but I am actually (and perhaps unwillingly) contributing to their permanent possession of my money. As a result, when we are deceived, we cannot assent nor choose to contribute to an end; instead, we are seen as a mere means to an end. In lying to or deceiving anyone, we are violating the Formula of Humanity under Kant's theory.

In regards to the “murderer at the door scenario,” the murderer is attempting to deceive us, according to Korsgaard's argument, so we can say that the murderer is

violating the Kingdom of Ends by seeing us as a mere means to their evil end. However, there is no argument from Kant that says once someone treats us as a means to an end that we are allowed to do the same. If we lie to the murderer when he asks us if we know the whereabouts of his victim, we too violate the Kingdom of Ends. First, the murderer cannot assent to the lie. A lie regarding your knowledge of their whereabouts, such as “I don’t know where they are,” cannot be taken to be a true statement. Even if the murderer somehow knew we were lying, he still cannot assent to the statement “I don’t know where they are” because it simply isn’t true. Also, the murderer cannot contain the end in himself. In lying to the murderer, our end is to prevent the murder, and we have not given the murderer a choice to contribute to this new end of ours. We are only giving the murderer the choice to contribute to our desired end if, for instance, we tell the truth and perhaps try to reason with him or get his mercy. However, we are not giving the murderer the choice to contribute to our desired end by giving him false information. We are now treating him as a means to our end. Again, even if the murderer somehow knew we were lying *and* he decided to not murder our friend, our action itself still treats him as a means to an end, so he would not be able to fully contain the end in himself.

Finally, under the lens of the Kingdom of Ends, to be used as a mere means to an end is to see someone as a tool as opposed to the rational being that they are. Lying, in particular, treats “someone’s reason as a tool” because we are choosing to withhold or manipulate information (1986 334). This hinders the other person’s ability to use their rationality, so we are not respecting the person’s status as a rational being by manipulating their rationality to achieve our ends. Korsgaard concludes that “this means that there cannot be a good reason for taking a decision out of someone else’s hands”

(1986 336). If the manipulation of any information is equated with the disrespect of someone as a person, it is no wonder why Kant would not want to say that lying can ever be the right thing to do.

Despite the conflicting outcomes produced by the Formula of Universal Law compared to the Formula of Humanity and Kingdom of Ends, Korsgaard attempts to reconcile them so that they can all make sense together to form the CI. In her reconciliation, she proposes her ideal and non-ideal theory. There is an ideal world where everyone acts according to the Formulas of Humanity and the Kingdom of Ends, and the non-ideal world, which is the world we inhabit, where not everyone acts according to the ideals. Our discussions so far have given us an idea of how to act in ideal circumstances, but in the non-ideal world where we are faced with the wrongdoing of others, Korsgaard argues that what is permissible changes in non-ideal circumstances.

“When dealing with evil circumstances we may depart from this ideal. In such cases, we can say that the Formula of Humanity is inapplicable because it is not designed for use when dealing with evil. But it can still guide our conduct. It defines the goal towards which we are working, and if we can generate priority rules we will know which features of it are most important. It gives us guidance about which of the measures we may take is the least objectionable.” (1986 347)

While Korsgaard does not explicitly say this, I believe the idea of using the Formula of Humanity to strive towards an ideal can be used to explain why we would lie to the

murderer as opposed to killing the murderer. Korsgaard does not elaborate on what she means by “priority rules” and how we can generate them using the Formula of Humanity. Nonetheless, we can recognize that if lying and killing can both prevent the murder of our friend, the least objectionable act would be lying.³

Her ideal and non-ideal theory seems appealing because it reconciles what seemed to be contradictory answers — our intuition that we should lie to the murderer and Kant’s response saying otherwise. By painting Kant’s response as to be upholding the ideal situation while permitting a different response under the non-ideal situation that is our reality, Korsgaard provides a charitable interpretation of Kant against rigorism. In her proposal, we still recognize the unconditionality of the moral principles in ideal circumstances, but we can also allow for exceptions in non-ideal circumstances. Exceptions can be made while still preserving Kantian ethics.

³ For more explanation as to why lying seems to be the least objectionable act, refer back to the introduction of the paper where I discuss the intuitions behind the moral weight of lying vs murder.

Chapter 3

Schapiro's Constitutive Approach

In Tamar Schapiro's essay "Kantian Rigorism and Mitigating Circumstances" (2006), she presents her own solution to rigorism called the constitutive approach. Before she presents her approach, she considers defenses of Kantian rigorism proposed by other philosophers, including Korsgaard's. Schapiro acknowledges some appealing aspects of Korsgaard's paper but argues that Korsgaard's argument faces some challenges. In particular, Korsgaard's is weakened by a note she discreetly hides in the fifth footnote of her paper. She admits that if the murderer were to somehow announce his intentions, and thus is not trying to deceive us, then we are not permitted to lie. This notion that we can only lie to deceivers is especially problematic in consideration of the real-life example of Nazis hunting Jews in WW2. The Nazis didn't have much reason to try to deceive anyone when they showed up to people's homes asking and inspecting for any Jews that were in hiding. They were government officials who were carrying out the orders of the government. In this case, Korsgaard would say that it is not morally permissible to lie to the Nazi and that we must share the whereabouts of those hiding in your home. In trying to preserve Kant's accounts in the *Groundwork* and "On the Supposed Right to Lie," Korsgaard puts too much focus on the details of a particular version of the "murderer at the door" case, but the "murderer at the door" scenario is meant to represent the much broader challenge of rigorism that Kantian ethics faces.

Despite the challenges Korsgaard's theory has, Schapiro still develops her approach by working off of some of Korsgaard's arguments. In her paper, Korsgaard

argues that deception is wrong because of its manipulative character. If deception constitutes manipulation, Schapiro further argues that deception, then, must be inherently wrong, or *wrong in itself*. If deception is wrong in itself, then no exceptions can make it right in any circumstances. This leads us to the root of the problem of rigorism — “actions required and prohibited by moral rules are right and wrong intrinsically” (2006 35). Because of rigorism, it may seem like deception could never be permissible, even in circumstances like the “murderer at the door” case. Schapiro notes that while Korsgaard demonstrates in her paper how deception can be seen as wrong in itself through her discussions of the Formula of Humanity and the Kingdom of Ends, Korsgaard also discusses how the Formula of Universal law says otherwise — i.e., deception is permissible under the Formula of Universal Law. Schapiro argues that Korsgaard does not explicitly reconcile these two opposing outcomes from the CI. In Korsgaard's theory, deception is permissible given non-ideal circumstances, such as the “murderer at the door” case. However, she does not explain how deception could be permissible even though, as she recognizes by her characterization of deception as manipulation, deception is wrong in itself. As a result, Schapiro seeks to explain how we can make the exception for lying in the “murderer at the door” case while still preserving the fact that deception is wrong in itself.

Before she begins her analysis of the “murderer at the door” case, Schapiro first examines a standard argument Kantians may give for a case in which deception is permissible — paternalistic deception. Although paternalistic deception is different from the deception in our case of interest, Schapiro uses it to outline the problems Kantians face with trying to permit deception. These are the same problems that she strives to

answer in the solution she proposes. According to the argument for paternalistic deception, deception is permissible when the one being deceived is not in the position to make decisions that count as their own. Paternalistic deception is often used to justify deceiving children and the mentally disabled because their mental capacities are governed by something else other than reason. For a child, they are not fully developed so their capacities are governed more by intuitions than reason. For the mentally disabled, their mental capacities are governed by a mental illness.

As discussed in the last chapter on Korsgaard, deception hinders another person's freedom to fully exercise their capabilities as a rational being. Deception treats the other person's rationality as a tool, and the agent can no longer contribute to the outcome of events, or an end; instead, their rationality is seen as a means to an end. We can understand this problem as deception interfering with another's autonomy. When we say an agent has autonomy, they have the freedom to make their own choices. By making it so an agent does not have the freedom to choose to contribute to an end, then deception is an interference of one's autonomy. Now going back to the problem of paternalistic deception, the motivation for paternalistic deception to begin with is that children and the mentally ill cannot make their own choices. Schapiro argues that if children and the mentally disabled don't have the freedom to make their own choices, then they do not have autonomy. If there is no autonomy to interfere with and deception is the interference of autonomy, then this is not a case of deception. When we do not tell the truth to someone who lacks autonomy, we cannot possibly be deceiving them. Not telling the truth in these cases would be performing acts of "intentional falsehood telling," as Schapiro phrases it, rather than acts of deception.

There are three problems that Schapiro identifies with the paternalistic deception argument. The first problem is that by looking at moral dilemmas, we are no longer looking at the domain of honesty and deception, but rather “intentional truth telling” and “intentional falsehood telling.” The latter domain is neutral in comparison to deception; however, we don’t want to say that in moral dilemmas, such as where there is a lack of autonomy, “intentional falsehood telling” is permissible. We still have the same intuitions — there should be constraints on “intentional truth telling” and “intentional falsehood telling” in the same way that we are seeking constraints on honesty versus deception. We need to find a bridge between the two domains here. This is one problem that Schapiro aims to address in the approach she proposes later.

The second problem is that the paternalistic argument does not make a distinction between an excuse and a justification. To justify an action is to say that an action is right; in contrast, to excuse an action is to still engage in the action while recognizing that the act is not morally right. For example, in our case of paternalistic deception, it would be impossible to justify it since deception is morally wrong in itself. If we were to try to justify paternalistic deception, the morally relevant description would not be of deception; instead, it would be a morally right act, such as promoting the development of a child. If we want to say that deception is permissible, we must first acknowledge the intrinsic wrongfulness of deception. In attempting to justify deception for paternalistic cases, the paternalistic deception argument fails to accept that deception is still morally wrong and can only be excused, not justified. Another way to understand the difference between an excuse and justification more intuitively is to see the question we are asking in moral dilemmas. Schapiro argues that we are not simply asking what is the right action

or if our acts are justified; instead, we are asking if we are permitted to depart from the standard to mitigate certain circumstances. In the “murderer at the door” scenario, we are not simply asking if lying to the murderer is justified; we are asking if we are permitted to not be honest due to the circumstances where telling the truth may lead to the murder of an innocent person.

If there was no difference between excuse and justification, Schapiro argues that there is then no difference between right thing to do and choosing the lesser of two evils. We should understand moral dilemmas like the “murderer at the door” case to be ones where we are not choosing the right action but rather choosing between the lesser of two evils. There cannot be a right action because both of our two choices of action — lie or assist a murderer — are wrong in themselves, and thus, cannot be a right action. Furthermore, when the agent chooses to lie, the agent “recognizes the choice worthiness of deception under the circumstances and regrets the fact that circumstances have made deception count as the choiceworthy alternative” (2006 42). The agent can recognize the wrongfulness of the action but also note that it was the most “choiceworthy” given that the circumstances made it so only wrongful actions could be committed. The integrity of the principles are preserved because it is in fact the “privation of the circumstances as opposed to a privation of the principles we have applied to those circumstances” that make it so we choose to deceive the murderer (2006 43).

The third problem is that paternalistic deception doesn’t extend to problems where deception seems warranted, like our “murderer at the door” case. In particular, the murderer does not lack autonomy. Unlike the argument for paternalistic deception, our goal is in a sense to interfere with the murderer’s ability to make a choice. We know that

in telling the truth, the murderer has the autonomy to proceed to find and murder the victim, so we intend to interfere with their decision making by manipulating the information we provide. Furthermore, the murderer must have autonomy if he is to be held responsible for the murder if we tell him the truth. As a result, the argument for paternalistic deception cannot be used to make deceiving the murderer permissible. The three problems with paternalistic deception all identify a problem we face to addressing the broader challenges with rigorism. In the rest of this chapter, I will present Schapiro's attempt to address the three problems she identifies and produce a more comprehensive response to rigorism.

Let us go back to the first problem she identifies. There are two domains — honesty vs deception and then “intentional truth telling” vs “intentional falsehood telling.” The reason for this divergence is because honesty and deception require there to be autonomy in the other person. Honesty demands the respect of that autonomy, and deception the interference of it. Schapiro argues that this suggests that autonomy is *constitutive* of honesty and deception. If there is a lack of autonomy, then honesty and deception cannot possibly be committed. For those that are not autonomous, the relevant actions would be more appropriately labeled as “intentional truth telling” and “intentional falsehood telling” as opposed to honesty or deception. Schapiro calls these *deficient*, or narrower, forms of honesty and deception. This is the general foundation of her *constitutive approach* — morally right and wrong actions constitute certain things, so if what is constitutive of the action is not present, then the morally right or wrong action in the fullest sense cannot be committed.

To understand more intuitively why autonomy would be constitutive of honesty and deception, let's consider what a respect for autonomy may mean in the overarching goals of being honest with someone. When we seek to be honest with someone, we are not simply looking to tell the truth. We respect their rational capacities, so in telling the truth, we aim to help them make decisions. In other words, when we want to act honestly, we are not focused on performing acts of truth-telling; instead, we have a goal of respecting and helping others. This is why honesty is different from and constitutes more than simple truth telling. We can understand truth telling as the "letter" of honesty since this may be the most simple definition that one gives for honesty. However, as we have discussed, this is not honesty in the fullest sense. There is also a "spirit" of honesty that captures the implicit goals we consider when we aim to act honestly, including respecting their rational capacities and helping them to make decisions. To act honestly in the fullest sense is to capture both the "letter" and the "spirit" of honesty.

In situations where we cannot be honest because there is a lack of something that is constitutive of honesty, Schapiro says that honesty has been "corrupted." We say here that honesty has been "corrupted" given that the circumstances make it so that honesty cannot be performed in its fullest sense. Nonetheless, we can still find a "surrogate" for honesty that can come "as close as possible to conforming to the ideal of human relations that enables acts of honesty in the first place" (2006 47-48). While honesty in the fullest sense is the "letter" and "spirit" of the act, the surrogate that approximates honesty aspires to capture the "spirit" of honesty because the spirit encapsulates the goals we aspire to achieve in being honest to begin with. Schapiro says that there are many actions that can be argued to capture the "spirit" of honesty, even deception. While Schapiro

doesn't explicitly explain how a deficient form of deception can capture the "spirit" of honesty, I conjecture that she believes that this would be the case if the act of deception helped maintain the sanctity and values of our social relations. In responding to this imperfect honesty with deception, we are actually using an imperfect version of deception as well. As aforementioned, the lack of autonomy that "corrupted" honesty is also a problem in pursuing deception because deception requires that autonomy. As a result, when Schapiro says that deception is permissible, we are excusing an imperfect or deficient form of deception. From this argument, Schapiro concludes that "if deception is wrong in itself yet sometimes permissible, this is because honesty, though intrinsically right, is nevertheless corruptible" (2006 48).

To address the third problem identified earlier (the paternalistic deception argument does not extend to the "murderer at the door" case), Schapiro argues that there is another thing that is constitutive of honesty — reciprocity. By deceiving, I prevent us from making choices that count as ours from a shared standpoint. Deception, from this view, can be understood as the refusal to reciprocate in achieving shared thoughts and actions. This deception as refusal concept presupposes the idea that morality includes joint participation in shared activities. We saw this idea of joint participation in Kant's Kingdom of Ends. Similar to autonomy, the notion that the refusal of reciprocation is constitutive of deception, we can understand honesty as the opposite — the acceptance of reciprocation. In the murderer at the door case, honesty is not possible because the murderer has refused to reciprocate in the joint participation in shared activities in his endeavor to take someone's life. In Kantian terms, the murderer has refused to participate in the Kingdom of Ends. Because the murderer has made it impossible to be honest in the

fullest sense in his refusal to reciprocate, deception is permissible if it is done in the spirit of honesty.

For those who do not reciprocate in the Kingdom of Ends, the question arises as to how we are to treat them and what obligations we have to them. Regarding actions that are permissible, Schapiro makes a distinction between an offense and a betrayal made to the Kingdom of Ends. When we consider in the normative sense when someone has offended us or our relationship with them, it is in a sense where they may not be behaving perfectly. The example she gives is a friend who is not a good listener. While they are not behaving well as a friend, we still recognize their commitment to the relationship. A betrayal, on the other hand, is where someone whose actions call into question their commitment to the relationship. This could be a friend who intentionally spreads rumors to ruin your reputation. By making this distinction, Schapiro hopes to capture the intuitions we have for moral dilemmas. For example, for paternalistic deception, we do not want to punish children and the mentally disabled for their inability to comply with the Kingdom of Ends, so we want to distinguish the lack of compliance with an inability, as opposed to a refusal. In contrast, the murderer has shown a refusal to comply with the Kingdom of Ends. Because the murderer has revealed that they lack a commitment to the moral relationship, Schapiro believes that the murderer has betrayed the Kingdom of Ends, and this betrayal mitigates our obligations to him. In particular, Schapiro says that we are “in a position to engage in constrained forms of manipulation” (2006 54).

Schapiro is careful to say our actions must be constrained as to not fall victim to the slippery slope challenge. If the murderer’s betrayal to the moral community and its ideals allows for mitigated obligations to him, there may be a concern of what obligations

we have to him, if any. Our standard moral principles and prohibitions don't apply to him, or at least, we are now allowed to make exceptions to the moral prohibitions. What, then, are we allowed to do to the murderer? Schapiro does not want her argument to be interpreted as endorsing any treatment towards the murderer, including extreme acts of torturing or killing the murderer. In addressing this, she writes,

“The murderer has not, through his betrayal, put himself
 “beyond the pale” of the moral community. By betraying
 the moral relationship, he has not thereby exited it” (2006
 54).

The betrayal of the Kingdom of Ends does not indicate an exit from the moral community since participation is not optional. Schapiro writes in a footnote that while she doesn't go into detail about what participation in the moral community entails, she believes it's along the lines of agency. If one has agency and can be held morally responsible for their actions, then they are part of the moral community, or Kingdom of Ends. The murderer, although he has betrayed the community, has an undeniable sense of agency that cannot be eliminated unless he were to suddenly lose his rational capabilities. If he has not relieved himself of the moral community and his moral obligations, then we are not relieved of the moral obligations we have to him. As a result, we are still obligated to seek the least objectionable act. If deception and murder are both sufficiently effective in derailing the murder, then we choose the act that mitigates the negative effects as to use honesty, in an aspirational sense.

Chapter 4

Cholbi's "Anything Goes" Objection to Schapiro

Michael Cholbi responds to Schapiro's constitutive approach in his paper "The Constitutive Approach to Kantian Rigorism." Cholbi appreciates how Schapiro provided a solution to the problems she sought to address — namely, the problem of permissible deception with the fact the deception is intrinsically wrong. He also appreciates how Schapiro avoided identifying one moral requirement as more significant or fundamental than another (i.e., that murder is worse than lying). This helps to make the Kantian argument, and Schapiro's theory involving Kantian ethics, more coherent.

However, Cholbi argues that it is not clear how the constitutive approach leads us to conclude that lying is the best action to take. Why can we not bribe, physically assault, or kill⁴ the murderer? Cholbi points out that if we excused a "corrupted" form of deception, what is to stop us from using the "corrupted" forms of other actions, such as threatening or killing. If we kill the murderer, Cholbi argues that we can say we do not wrongfully kill him because his actions "made it impossible for me to refrain from murdering him in the spirit proper to not murdering" (2012 444). In other words, just as autonomy and reciprocity constitute honesty, autonomy and reciprocity can be argued to constitute the act of not killing someone. Most people would not understand killing a person to be equal to the killing of an animal or a tree, nor would people place the same moral judgment on a person killing another person compared to an animal or object

⁴ While Cholbi says "murdering the murderer," I'm intentionally choosing to say "kill" as opposed to "murder" since "murder" can be understood in a legal sense as killing with malicious intent. "Killing" here should be taken neutrally with no implications or assumptions of intent or context; it is simply the act of taking someone's life.

killing a person. This suggests that killing, and *not killing*, someone has elements of respect for others' autonomy in a reciprocal manner that is natural of the Kingdom of Ends, as discussed in Schapiro's paper. As a result, because the act of "not killing" has been corrupted in the same way that honesty has been corrupted, it may seem as though a corrupted version of killing is permissible in the same way that corrupted deception is permissible. This serves as the motivation and reasoning behind Cholbi's challenge that Schapiro's argument is too permissive since once we recognize that we are unable to perform our duties in the fullest sense, it seems like we can excuse any corrupted form of an immoral act. This is the "anything goes" objection — once the letter of a morally permissible act is compromised, any act is permissible.

Even if one were to argue that killing the murderer is permissible, it cannot be that any action is permissible, as argued in the "anything goes" objection. Not torturing someone can be argued as constituting autonomy and reciprocity along similar lines to not killing someone. In that case, are we now permitted to senselessly torture the would-be murderer to prevent the murder? This seems to go against our intuitions of morality. Even if someone commits a moral wrongdoing, there should still be moral limitations on how we respond and treat the other person. Schapiro herself recognizes this too. As discussed in our examination of Schapiro's constitutive approach, Schapiro tries to prevent the "anything goes" objection by making a distinction between a betrayal and an exit from the moral community. Even though Schapiro said that the murderer has not exited the moral community and thus has not lost his status as a moral subject, Cholbi thinks that the murderer's betrayal of the Kingdom of Ends itself warrants the "anything goes" objection.

Schapiro says that the murderer betrayed the ideal of the Kingdom of Ends, so Cholbi interprets that this is a betrayal of all the norms of the Kingdom of Ends. Honesty is impossible given that the murderer has not respected the ideal of reciprocity, but in the Kingdom of Ends, there is always a sense of reciprocity since everyone respects one another as a rational being and aims to promote others' ends in addition to their own. To show a lack of commitment to the ideal of the Kingdom of Ends, this is not just a betrayal to the integrity of honesty but to the Kingdom of Ends as a whole. If this is a betrayal of all norms, it seems that everything is out the window.

If the murderer has chosen to not respect nor participate in the Kingdom of Ends, it is unclear why the norms and obligations from the Kingdom of Ends should still apply to the murderer. Cholbi argues that there is nothing in the constitutive approach telling us that we must not kill the murderer because that would be morally wrong. As aforementioned, the constitutive approach permits us to use "corrupted" forms of deception, and it could be argued that the constitutive approach does not prohibit us from using the corrupted forms of other acts, such as killing. It may be morally admirable to lie to someone as oppose to murder, but the issue is that nothing in the constitutive approach prohibits murder. Perhaps one may see no issue in killing the murderer in this case. Nonetheless, Cholbi is trying to raise the moral general criticism that the constitutive approach does not seem to set any moral prohibitions for moral dilemmas. If someone betrays the ideal of the Kingdom of Ends, then we can mistreat them for no good reason (e.g., torturing them), and there seems to be no actions that are strictly prohibited.

Cholbi provides a brief outline for an alternative to the constitutive approach that he calls the "fine print" approach. Cholbi's approach is focused on the Formula of

Universal Law from the CI. His solution is to embed a conditional antecedent in the maxim that you are trying to universalize. For our “murderer at the door” case, rather than having the maxim simply permit deception (since this would fail Kant’s universalization), adding the conditional antecedent “If deception would prevent the death of an innocent person” does not fail the universalization test. As a result, Cholbi argues that a defensible exception can be provided when given the context because this would pass the Formula of Universal Law.

Cholbi’s approach, however, seems to bring us back to square one. In her paper, Korsgaard argues that while deceiving the murderer passes the Formula of Universal Law, Kant says that lying is not permitted because it would undermine the other two parts of the CI — the Formula of Humanity and the Kingdom of Ends. As a result, this strategy dismisses two crucial parts of Kant’s theory. However, the goal I’ve set out in this paper is to provide a defense of Kantian ethics because the strengths of his theory make me hesitate to dismiss the theory for its supposed rigorism. In the final chapter, I will present a possible defense for Schapiro against the “anything goes” objection.

Chapter 5

Defense of Schapiro

Cholbi argues that a betrayal of the Kingdom of Ends nullifies our obligations to the murderer. Let us remind ourselves of the problem between a betrayal versus an exit. Schapiro argues that we still have moral obligations to someone who betrays the Kingdom of Ends; we don't have any moral obligations for those who have exited or are not part of the Kingdom of Ends. Cholbi believes that we actually do not have moral obligations to those who betray the Kingdom of Ends because of the constitutive approach. Cholbi does not explicitly say whether he believes that betraying the Kingdom of Ends is the same as exiting the Kingdom of Ends. However, we should note that Schapiro notes how she believes exiting the moral community is actually impossible because participation in the Kingdom of Ends is involuntary. Schapiro sets a high bar for an exit from the Kingdom of Ends — one must no longer have a sense of agency to truly exit the moral community, but agency is something that one arguably cannot simply get rid of. A murderer certainly does not get rid of his agency by betraying the Kingdom of Ends — he still continues to act as a moral agent so long as he is rational and autonomous. As a result, even with granting Cholbi's challenges against Schapiro, we still ought to say that the murderer is part of the Kingdom of Ends. Schapiro argues that by being a part of the moral community, we still have moral obligations to the murderer, but Cholbi points out that moral obligations seem impossible under the constitutive approach. How do we reconcile this issue?

I believe that the key is to recognize is that we have broader obligations to the moral community as a whole in addition to the obligations we have to other people. Specifically, we have the obligation to try to encourage others to be good members of the moral community, or at the very least, we should not discourage them. For example, just as we are held responsible if we harm others, we should also be responsible if we encourage others to commit harm even if we do not commit the harm ourselves. To encourage wrongdoing in others is just as bad as committing the wrongdoing ourselves even if we are not the agents committing the actions — we are knowingly and intentionally encouraging others to commit morally wrong actions. We shouldn't just stop here, though. We should also encourage others by telling them what it means to respect and help others. For example, when we tell someone that how they deal with conflict is aggressive and insensitive, we want to follow up by telling them how they can be change moving forward and be better. That way, the conversation is productive as opposed to simply criticizing and punishing them for their actions. This will help them be more aware of how they deal with conflict moving forward, and this can be beneficial to us as well in that they will manage conflicts more appropriately. As a result, we have an obligation to ourselves and others to encourage others to be good members of the moral community.

We can now phrase this under Kantian terms to say that we have the obligation to promote the Kingdom of Ends. While we are not responsible for the murderer betraying the ideals of the Kingdom of Ends, we should still act in a way that encourages the murderer to eventually return to the ideals. In other words, while we are not responsible for other people's wrongdoing, we are responsible for our actions in limiting or

encouraging other people's abilities to be a part of the Kingdom of Ends. Even though the murderer has betrayed the ideals of the Kingdom of Ends in attempting to murder someone, the murderer still has the capability of eventually respecting and upholding the ideals. Because there is this capability, we should act in a way that acknowledges and respects this capability. In responding to the murderer at the door, we should lie as opposed to kill them because killing them permanently hinders the murderer from ever returning to the ideals of the Kingdom of Ends. Torture is less permanent than killing the murderer; however, the mental or physical trauma sustained by the murderer can serve as difficult obstacles that the murderer has to overcome. Moreover, torture may greatly discourage the murderer from wanting to return to the Kingdom of Ends as it may invoke anger, resentment, and general pessimism towards others. Deception is temporary, and the murderer may never realize that they were deceived. As a result, while it may initially seem that the constitutive approach allows for "corrupted" killing or "corrupted" torture, we are obligated to only use corrupted deception since this is the act that least hinders the murderer's eventual return to the ideals of the Kingdom of Ends.

This defense of Schapiro aligns with other arguments that Schapiro has developed in regards to Kantian ethics. In her "What Is a Child?" essay, Schapiro discusses how we should treat children in the context of the Kingdom of Ends. There is a predicament with children in that despite their status as undeveloped autonomous, rational agents, they are still seen as citizens in the Kingdom of Ends. We are reminded of Rawls's nonideal versus ideal theory that Korsgaard used as a framework for her theory. In this case, a child's dependence can be seen as a nonideal status, and the developed rationality and autonomy in adults is the ideal status. If a child's dependence is considered nonideal,

Schapiro sees the “the undeveloped nature of children’s agency as an obstacle to morality, a condition which in principle ought to be eliminated” (1999 735). Children’s dependence is an obstacle to morality because they are not fully developed and responsible moral agents. As adults, we must recognize how difficult it is to be an undeveloped agent. “Accordingly, we should make it our end to do what is in our power as adults to help children work their way out of childhood” (1999 735). We must refrain from being too paternalistic to a child that it hinders their development as autonomous beings. It is not to say that we treat them as adults, but rather that we should start modeling behavior and encouraging them to develop their autonomy.

While there are obvious differences between children and murderers, our obligations to both of them are the same. Murderers also pose as an obstacle to morality. They are agents who are capable of upholding the ideals of the Kingdom of Ends, yet they have failed to do so. Nonetheless, they are still members of the Kingdom of Ends since all those with agency are participants. We all have the obligation to promote the Kingdom of Ends, and thus, we have the obligation to not act in a way that would hinder the murderer’s ability to realize their capabilities. We can frame our obligation as the following — For those members who are in the Kingdom of Ends but not upholding the ideals of the Kingdom of Ends for some particular reason, we have the obligation of recognizing their capabilities to eventually uphold the ideals and promoting (or, at the very least, not hindering) their ability to do so.

Further Considerations

My goal in this paper was to provide a thoroughly developed defense for Kantian ethics against rigorism, specifically under the examination of the “murderer at the door” case. There are still some questions, however, that could be asked about the murderer at the door case. In the first chapter of this paper, I drew upon our intuitions about lying vs killing to help us begin to understand the problem raised in the “murderer at the door” case. Without considering what Kantian ethics tells us, our natural solution to the situation is to lie because in considering lying vs murder (or abetting in a murder), lying seems less bad. While I have presented a defense for Kantian ethics to help us arrive at the conclusion that we should lie in the “murderer at the door” case, the question remains as to whether my proposed solution aligns with or is a good substitute for the initial intuitions we had about the case. Does Kantian ethics align with our moral intuitions, bring clarity to them, or go against them?

One may argue that lying and killing should not fundamentally be wrong in the same way. The argument can be made that lying is not as bad as murder, so there should be a hierarchy of moral rights and moral wrongs. In this paper, I only argued that we should lie as opposed to murder the murderer to promote the Kingdom of Ends. I still maintained the argument that lying is wrong in itself the same way that murder is wrong in itself. I, nor did those whose works I discussed, explicitly argue for murder being a more morally wrong action than lying; instead, we sought to find ways to explain how we are morally obligated to lie, as opposed to kill or assist in a killing, in the “murderer at the door” case. However, one may argue that the intuitions behind these arguments is that lying seems to be a better solution than murder. If we disregarded the context behind the

“murderer at the door” case, could it be argued that killing is intrinsically morally worse than lying? If so, then should we seek out a ranking of moral rights and moral wrongs?

The need for a ranking of moral rights and moral wrongs is not self-evident and goes beyond the scope of this paper. Nonetheless, I believe that a ranking in a moral theory is an interesting point to consider further. The first question to ask is whether it is important for a moral theory to have rankings. If so, then we should consider what this means for Kantian ethics. Is Kant’s theory less plausible if it does not have a ranking system? While this requires much more in-depth exploration and discussion, my initial feelings towards a ranking system is that it could be very complicated for a moral theory to implement and Kantian ethics, as shown in this paper, can conclusively guide us to the right action. Just as we explored in this paper, I can see how circumstances ask us to make exceptions to the rankings. We may find forms of lying worse than forms of killing. For example, infidelity is worse than killing someone out of self-defense. To make rankings out of specific moral rights and wrongs seems quite complicated, and one would have to find a common thread throughout all the rankings to provide a theory behind the rankings.

I do not, however, believe that a ranking system would be impossible. While it did not initially seem coherent with Kantian ethics, as I argued in the last chapter, some actions are preferable to others if they meet our broader obligations to the Kingdom of Ends. The ranking could be that of preferability, as opposed to inherent wrongness, to maintain coherence with Kantian ethics. Further consideration, while complementing this paper, can be given for how a ranking system could work with Kantian ethics, if deemed necessary.

For now, I've laid out the argument of how Schapiros' constitutive approach can help us conclude that deception is permissible in the "murderer at the door" case using Kantian ethics. One may find this argument unsatisfactory — that deception is not only permissible, but the right thing to do. In this case, this seems to directly contradict Kant's theory, so to make the argument that it is the right thing to do, one would have to develop the argument with another ethical theory. The intention of my paper was to defend Kant's argument against rigorism, which I've done, so I won't pursue an alternative to Kant's theory to accommodate this intuition some may have.

Bibliography

Works by Kant

Groundwork of the Metaphysics of Morals, 1785. Trans. and ed. Mary Gregor and Jens Timmermann. Cambridge: Cambridge University Press, 2011.

“On the Supposed Right to Lie.” *Critique of Practical Reason and Other Works on the Theory of Ethics*, 1797. 4th ed. Trans. Thomas Kingsmill Abbott. London: Kongmans, Green and Co., 1889.

Lectures on Ethics. 1775-1780. Trans. Louis Infield. Indianapolis, Hackett Press, 1980.

Secondary Sources

Cholbi, Michael. “The Constitutive Approach to Kantian Rigorism.” *Ethical Theory and Moral Practice*, vol. 16, no. 3, 2012.

Korsgaard, Christine. “The Right to Lie: Kant on Dealing with Evil.” *Philosophy and Public Affairs* 15, no. 4, 1986.

Korsgaard, Christine. *Creating the Kingdom of Ends*. Cambridge University Press, 2004.

Schapiro, Tamar. “What Is a Child?” *Ethics*, vol. 109, no. 4, 1999.

Schapiro, Tamar. “Kantian Rigorism and Mitigating Circumstances.” *Ethics*, vol. 117, no. 1, Oct. 2006.