

FREEDOM AND THE UNINTELLIGIBILITY OF EVIL
IN KANT'S MORAL PHILOSOPHY

BERKE AYDOĞAN

BOĞAZİÇİ UNIVERSITY

2025

FREEDOM AND THE UNINTELLIGIBILITY OF EVIL
IN KANT'S MORAL PHILOSOPHY

Thesis submitted to the
Institute for Graduate Studies in Social Sciences
in partial fulfillment of the requirements for the degree of

Master of Arts
in
Philosophy

by
Berke Aydoğan

Boğaziçi University

2025

Freedom and the Unintelligibility of Evil in Kant's Moral Philosophy

The thesis of Berke Aydoğan

has been approved by:

Prof. Berna Kılınç
(Thesis Advisor)

Prof. Chrysi Sidiropoulou

Prof. Aliye Kovanlıkaya
(External Member)

January 2025

DECLARATION OF ORIGINALITY

I, Berke Aydoğan, certify that

- I am the sole author of this thesis and that I have fully acknowledged and documented in my thesis all sources of ideas and words, including digital resources, which have been produced or published by another person or institution;
- this thesis contains no material that has been submitted or accepted for a degree or diploma in any other educational institution;
- this is a true copy of the thesis approved by my advisor and thesis committee at Boğaziçi University, including final revisions required by them.

Signature.....

Date

ABSTRACT

Freedom and the Unintelligibility of Evil in Kant's Moral Philosophy

In this thesis, I examine a classical objection to Kant's moral philosophy. Kant argues that the law of free will is the moral law. Carl Leonard Reinhold responded with the criticism that free and immoral action is impossible according to Kant's view. In the first chapter of the thesis, I examine Reinhold's criticism and Kant's reply to Reinhold. In the second and third chapters, I present Kantian responses to Reinhold in contemporary philosophy. In the last chapter, I offer a reading of Kant's posthumously published notes in the context of his debate with Reinhold. I argue that Kant's view suggest that evil cannot be explained in terms of freedom, but as a fact about the phenomenal world.

ÖZET

Kant'ın Ahlak Felsefesinde Özgürlük ve Kötülüğün Anlaşılabilirliği

Bu tezde, Kant'ın ahlak felsefesine verilen klasik bir eleştiri incelenmektedir. Kant'a göre, özgür iradenin yasası ahlak yasasıdır. Carl Leonard Reinhold ise Kant'ın görüşünün, özgür ve kötü eylemi imkansız hale getirdiğini savunmuştur. Tezin ilk bölümünde Reinhold'un eleştirisi ve Kant'ın Reinhold'a verdiği cevap incelenmektedir. İkinci ve üçüncü bölümlerinde ise çağdaş felsefede Reinhold'un itirazına verilen Kantçı cevaplar incelenmektedir. Tezin son bölümünde ise Kant'ın ölümünden sonra yayınlanan notları Reinhold ile olan tartışması bağlamında değerlendirilecektir. Tezde sunulan okumaya göre, Kant için kötülük, özgürlük çerçevesinde değil yalnızca fenomenler dünyasına ait bir olgu olarak açıklanabilir.

ACKNOWLEDGMENTS

I would like to thank my advisor, Berna Kılınç, for accepting me as her student and for endlessly supporting me in writing this thesis. Our discussions and her comments on the drafts of my thesis have been invaluable for me. I also thank Chryssi Sidiropoulou for her continuous support since my undergraduate years. She always took the time to listen to me compassionately and encouraged me in my academic journey. A special thanks to Lucas Thorpe, whose courses helped me discover my interest in Kant's philosophy. Without his courses and insightful discussions, I could not have raised the question that I address in this thesis. I also thank Andrew Fyfe for his unforgettable friendship and academic support, and Aliye Kovanlıkaya for serving on my committee.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION	1
CHAPTER 2: FREEDOM AND THE PROBLEM OF EVIL	7
2.1 Kant's exposition of the law of free will	8
2.2 Ulrich, Schmid, and Reinhold.	12
2.3 Kant's response to Reinhold.....	16
CHAPTER 3: TWO POPULAR RESPONSES.....	20
3.1 The capacity reading.....	20
3.2 The <i>Wille-Willkür</i> reading	23
CHAPTER 4: KORSGAARD'S RESPONSE: THE UNINTELLIGIBILITY OF EVIL.....	29
4.1 Noumenon and phenomenon	29
4.2 Korsgaard's <i>Groundwork</i> reading	32
4.3 Freedom as a postulate.....	38
CHAPTER 5: THE GOOD HUMAN WILL	45
5.1 Kant's notes and <i>Wille-Willkür</i>	47
5.2 Goodness of human will and evil maxims	50
CHAPTER 6: CONCLUSION.....	59
REFERENCES.....	62

CHAPTER 1

INTRODUCTION

The most distinctive feature of Immanuel Kant's moral philosophy is its promise for grounding morality in freedom. While Kant believes that we cannot know whether we are free in a theoretical sense, he argues that the idea of freedom is a necessary and inevitable part of our practical and moral life. He writes in *Groundwork of the Metaphysics of Morals* "a free will and a will under moral laws are one and the same" (4:447).¹ Kant's view is termed "the reciprocity thesis" by Henry Allison (1990).

However, Kant faces a classic objection. If free will and a will under the moral law are the same thing, how is free and immoral action possible? It is this objection that I hope to answer in the present work. Carl Leonard Reinhold's version of the objection will be the focus of this present work. Reinhold famously argued that if the will is free insofar as it is governed by the moral law, then the human will cannot be "impure", or evil, and free at the same time (2021, p.268).

Reinhold initially thought that this conclusion did not follow from Kant's argument, but from the misappropriations of Kant's philosophy during Kant's era. He offered a clarification and wrote that Kant meant freedom to consist in the capacity to make a choice for or against the moral law. However, Kant responded to him in the *Metaphysics of Morals* that "freedom can never be located in a rational subject's being able to choose in opposition to his lawgiving reason..." (6:226).²

¹ *Groundwork of the Metaphysics of Morals*, trans. by Mary Gregor, in *Practical Philosophy*, ed. by Mary Gregor, Cambridge: Cambridge University Press, 1996, pp. 37–108. Abbreviated in text as G.

² *The Metaphysics of Morals*, trans. by Mary Gregor, in *Practical Philosophy*, ed. by Mary Gregor, Cambridge: Cambridge University Press, 1996, pp. 353–604. Abbreviated in text as MM.

Reinhold, being surprised by Kant's answer, maintained his own position and concluded that Kant's moral philosophy is untenable.

In his reply to Reinhold, Kant did not totally reject the reality of freely chosen evil. He wrote that experience shows that "the human being" has the ability (*Vermögen*) to choose for and against the moral law. However, he rejected that what is observable through experience, or "appearances" can define or explain (*erklären*) a "supersensible object", such as "free choice (*Willkür*)" (Ibid.).³ The second chapter of this thesis is an examination of Reinhold's objection and Kant's reply to Reinhold.

The third chapter concerns two popular contemporary responses to Reinhold's objection. The first one is known as the capacity reading.⁴ According to this reading, Kant's identification between free will and the moral law must not be understood as a strict reciprocity, but a loose one. It has been argued that rather than the actual exercise of morality, Kant identified the capacity for the exercise of morality with freedom. Whether we act from the moral law or not, we are always free; just as an eye without vision is still an eye, but one that does not exercise its capacity for vision (2023, p.42). However appealing this reading may be, there is an abundance of textual evidence that suggests that Kant meant the actual exercise of moral capacity as what defines us as free rather than mere capacity.

The other popular response is known as the *Wille-Willkür* response. The proponents of this response argue that Kant became aware of the need for a distinction that could accommodate free evil in his later works due to Reinhold's objection. In the *Metaphysics of Morals*, Kant defines *Wille*, or the will, as the

³ *Erklärung* in Kant's usage is translated as explanation, definition, explication, exposition, and declaration. In this thesis, I mainly use "explanation" and "definition" depending on the translation cited.

⁴ This reading is defended by H.J. Paton (1948) and Jochen Bojanowski (2023)

legislative capacity responsible for lawgiving (6:226). He defines *Willkür*, which is usually translated as the power of choice, as the faculty from which maxims of actions arise. It has been argued that while *Wille* cannot be both free and evil, *Willkür* can be because freedom of *Willkür* concerns only maxims, not lawgiving. Thus, *Willkür* can either reject or conform to the laws given by *Wille* through its maxims and can remain a free *Willkür*. However, this response does not take into account the paragraph that follows Kant's distinction between *Wille* and *Willkür* in the *Metaphysics of Morals* where Kant responds to Reinhold. Kant rejects that we can define free *Willkür* as one capable of acting against the moral law (6:226).

The fourth chapter concerns one novel response to the objection by Christine Korsgaard (1996). Korsgaard argues that Kant's identification between freedom and a will under the moral law supports the idea that evil is unintelligible, not the idea that it is impossible (p.15). Korsgaard's Kant reading is based on the two distinct employments of reason. The theoretical employment of reason is explanatory. We can theoretically explain the world only as it appears to us or as phenomenon. We cannot theoretically explain the world as it is in itself or as noumenon because knowledge is restricted to experience. The practical employment of reason is action-guiding. It is not concerned with explaining the world, but with providing principles of action. For Korsgaard, freedom is a "practical" concept that belongs to the practical employment of reason (Ibid.).

For Korsgaard, practical deliberation requires that we "act under the idea of freedom" when we make decisions about what to do (p.5). While we cannot theoretically know whether we are free, practical deliberation requires that we act as if we are free. Korsgaard argues that Kant's views show only that we act according to the moral law in a practical sense when we think of ourselves as free. In

considering ourselves as free, we also consider ourselves as noumena acting in accordance with the moral law. For Korsgaard, evil is unintelligible because it does not make sense normatively or practically when we regard ourselves as free. However, the thought of freedom and the noumenal world has only a practical, not a theoretical function for Korsgaard. Theoretically, it is possible, though not knowable and intelligible, that noumena may not act according to the moral law, yet remain free. Korsgaard concludes that evil is unintelligible in a practical sense, but still we need to theoretically leave open the idea of free and immoral action on the noumenal level (p.25).

In the fifth chapter of the thesis, I present my reading of Kant's two-page loose leaves posthumously published under the title *Preliminary Notes and Reflections to the Introduction to the Metaphysics of Morals, Before 1797*.⁵ Kant's notes have not received sufficient attention in the context of Reinhold's criticism. One of my main goals is to show that Kant's notes help us better understand the source of his disagreement with Reinhold. While Kant previously wrote in *Metaphysics of Morals* that only the human being as phenomenon shows a capacity for evil, he did not elaborate on his view (6:226). I read Kant as providing a clarification in his notes that has gone largely unnoticed.

In *Preliminary Notes*, Kant writes that the human will (*Wille*) is good and we cannot explain evil "from the intelligible substrate of free will" (p.231). I argue that for Kant, the phenomenal fact that we sometimes act immorally does not indicate that we have an evil will, and thus this fact is insufficient for explaining or defining immoral actions as free.

⁵ *Preliminary Notes and Reflections to the Introduction to the Metaphysics of Morals, Before 1797*, in *Kant's Early Critics on Freedom of the Will*, ed. by J. Noller and J. Walsh, Cambridge: Cambridge University Press, 2022, pp. 229–231. There is no Akademie edition available in English.

After Kant rejects that we can explain evil in relation to free will, he also writes that we can explain evil “only mechanically” (ibid.). In my reading, this remark helps us consistently make sense of Kant’s position against Reinhold that only the human being as phenomenon shows us a capacity for choosing evil. Kant does not provide a clarification for what he means by mechanical explanation of actions. I suggest that it follows from Kant’s response that immoral actions cannot be explained in relation to human volition or freedom, but in terms of observable facts about the world and individuals. In other words, evil can be explained only as a fact about the phenomenal world.

In my reading, I benefit from the language of worlds as well in order to argue that there is no contradiction between having evil maxims and possessing a good will for Kant. The idea that the human will is part of the noumenal world is an important implicit premise in Kant’s response, which I believe helps us appreciate Kant’s response better. As Kant argues multiple times, the human volition never fails to be the source of the moral law (MM 6:226). In his notes, he adds that “Precisely for that reason it [*Wille*] is good in all human beings and there is no unlawful volition (2022, p.231)”. The fact that we sometimes have evil maxims, for Kant, does not show that the human will is the source of these maxims because we observe evil only when we “consider” ourselves “as phenomenon” (Ibid.). In his official response, Kant also rejects that “appearances” can define freedom. In a parallel way, I argue that what happens in the phenomenal world, namely that we act immorally, can explain the will or the noumenal world, of which our wills are a part. Thus, I conclude that the ultimate goodness of human volition, for Kant, does not contradict the reality of evil in the world.

The other purpose of the thesis is to discuss Kant's remark on the "intelligible substrate of free will" and mechanical explanation for an interpretation of the idea that evil is unintelligible. I argue that the explanation of why we act immorally cannot extend beyond the phenomenal world due to the limitations of our knowledge. As a result, why someone would act immorally cannot find an answer in the volition of the agent without considering the influence of the conditions of the world. I conclude that evil is unintelligible because we cannot explain evil maxims as arising from our will, but from the conditions of us humans and the world.



CHAPTER 2

FREEDOM AND THE PROBLEM OF EVIL

Kant argues that we are responsible for both good and evil. In *Religion within the Boundaries of Mere Reason*, Kant writes:

The human being must make or have made himself into whatever he is or should become in a moral sense, good or evil. These two [characters] must be an effect of his free power of choice, for otherwise they could not be imputed to him and, consequently, he could be neither morally good nor evil. (6:44)⁶

Unless the “characters” of the human being, which express good and evil principles, are “an effect of his free power of choice (*Willkür*)...they could not be imputed to him”. In the same work, Kant also argues that we have an innate propensity (*Hang*) to evil, which he also calls radical evil. He defines radical evil as the propensity of humans to subordinate the moral law to maxims of self-love (6:36).⁷ While Kant calls this act of subordination a propensity, he maintains that this, and the resulting immoral action, is imputable to our free choice.

Although Kant’s remarks in *Religion* show that he endorses the idea of freely chosen evil, his arguments in his other works were read by critics as excluding this idea. Kant’s argument in *Groundwork* that “a free will and a will under the moral laws are one and the same” is one of them (4:447). Henry Sidgwick (1888), for instance, charged Kant with the inconsistency of using two irreconcilable conceptions of freedom in different contexts (p.405). In Sidgwick’s understanding, one concept, as given in *Religion*, applies to all actions, whether good or evil. The

⁶ *Religion within the Boundaries of Mere Reason*, trans. by Allen Wood and George di Giovanni, in *Religion and Rational Theology*, ed. by Allen Wood and George di Giovanni, Cambridge: Cambridge University Press, 1996, pp. 39–216.

⁷ A maxim is a subjective principle of action in the form of “I will do Y in order to get X”. Kant believes that humans are not determined by desires, but they always act on the basis of maxims that incorporate these desires (*Religion* 6:24).

other concept is deployed by Kant for the justification of morality on the idea of freedom. Sidgwick argued that it is not possible to accept both concepts without giving full treatment to the relation between the two. He did not consider a reconciliation between two concepts impossible.⁸ However, he also believed that this would eventually undermine Kant's ambitious project of grounding morality in freedom, or Kant's "spirit-stirring appeal to the sentiment of Liberty" (p.412). In the next section, I present Kant's argument for the reciprocity between freedom and the moral law in *Groundwork*.

2.1 Kant's exposition of the law of free will

The first two sections of *Groundwork* present an analysis of ordinary moral concepts, such as good will and duty. He calls the method of the first two sections analytic. Kant does not go beyond the analysis of these concepts, and he does not answer why it is necessary for us to be moral in these two sections. *Groundwork III* attempts to provide this answer by first showing that the law of a free will is the moral law, and that we necessarily think of ourselves as free and subject to the moral law. In what follows, I present only the first step in his argument as the other step does not directly relate to the topic of this thesis. The argument is short, but dense. For this reason, I examine it in parts.

Will is a kind of causality of living beings insofar as they are rational, and freedom would be that property of such causality that it can be efficient independently of alien causes determining it, just as natural necessity is the property of the causality of all nonrational beings to be determined to activity by the influence of alien causes. (G 4:447)

⁸ Sidgwick (1888) called Kant's account of freedom in *Religion* neutral or moral freedom, while he called Kant's other account in his works, such as *Groundwork* and the *Critique of Practical Reason*, rational or good freedom (p.407).

The third section begins with a contrast between natural necessity and will. Kant calls the former “the causality of all nonrational beings to be determined into activity by the influence of alien causes”. By “nonrational beings”, Kant means nonhuman animals. In Kant’s classification, animals have a volitional capacity, but it is different from the one that humans possess (MM 6:213). Animal volition is determined by alien causes, such as instincts provided by nature, which expresses the idea of natural necessity for Kant.

Kant defines will as “a causality” of rational beings, and freedom as the “property” of that causality to be “efficient independently of alien causes determining it”. The will of a rational being operates differently from animal volition because it can self-determine independently of the influence of alien causes. Kant notes that his definition of freedom up to this point is only negative, and therefore “unfruitful for insight into its essence”. Freedom defined negatively is the idea of an efficient causality that is not determined by natural necessity. This definition is negative in the sense that it defines freedom not by what it is but by what it is not. In this respect, the negative concept of freedom is “unfruitful” regarding “its essence” because it does not tell how a being free from natural necessity acts. The positive concept of freedom defines freedom as to its essence and tells how a negatively free being acts.

The preceding definition of freedom is negative and therefore unfruitful for insight into its essence; but there flows from it a positive concept of freedom, which is so much richer and more fruitful. Since the concept of causality brings with it that of laws in accordance with which, by something that we call a cause, something else, namely an effect, must be posited, so freedom, although it is not a property of the will in accordance with natural laws, is not for that reason lawless but must instead be a causality in accordance with immutable laws but of a special kind; for otherwise a free will would be an absurdity. (Ibid.)

A lawful conception of freedom is necessary because Kant takes it for granted that “the concept of causality”, whether the causality of nature or will, involves the concept of law. In Kant’s conception, the necessary relation between a cause and an effect is a relation that we express in terms of laws (B124). Laws of nature express this necessity for events in the world. In *Religion*, for instance, Kant writes that the idea of a lawless will is an “absurdity”, just as the idea of a lawless natural necessity is an absurdity (6:35). If the will is a causality that can lead to effects, it must have a law as well. The positive concept of freedom, which specifies the law of a free will, “flows from” the negative concept upon a consideration of the law from which a will free in the negative sense would act:

Natural necessity was a heteronomy of efficient causes, since every effect was possible only in accordance with the law that something else determines the efficient cause to causality; what, then, can freedom of the will be other than autonomy, that is, the will's property of being a law to itself? But the proposition, the will is in all its actions a law to itself, indicates only the principle, to act on no other maxim than that which can also have as object itself as a universal law. This, however, is precisely the formula of the categorical imperative and is the principle of morality; hence a free will and a will under moral laws are one and the same. (4:447)

Kant calls natural necessity “a heteronomy (*Heteronomie*) of efficient causes”.

Heteronomy is the state of being governed by external influences, such as natural laws (4:433). Etymologically, the word heteronomy comes from the combination of two Greek words, *heteros* (other) and *nomos* (law). *Heteros-nomos* means being governed by the law of someone or something else. Autonomous, which comes from *autos* (self) and *nomos* (law) in Greek, means being governed by one’s own law. If no other law, as in the case of heteronomy, can determine the free will, then it is autonomous in the sense that the will is a “law to itself”. A free will finds its law in itself, and therefore acts on its own law. Without too much clarification, Kant writes that “the will’s property of being a law to itself” can be formulated as the principle

“to act on no other maxim than that which can also have as object itself as a universal law”. This is the first formulation of the categorical imperative, which Kant in the same work also formulates as “act only in accordance with that maxim through which at the same time you will that it become a universal law” (4:422). Kant concludes that “a free will and a will under moral laws are one and the same”.

Kant reiterates the same argument in the *Critique of Practical Reason* and similarly concludes that “Thus freedom and unconditional practical law reciprocally imply each other” (5:29).⁹ Kant’s view has been given different names. Henry Allison (1990) termed Kant’s view “the reciprocity thesis” (p.201). Though Kant himself never concluded that non-universalizable, or evil maxims, are not free, his account was taken to imply this conclusion and contradict his own account in *Religion*. However, I believe that the accusation of inconsistency between works is unfair because Kant also writes in *Religion* that “To think of oneself as a freely acting being, yet as exempted from the one law commensurate to such a being (the moral law), would amount to the thought of a cause operating without any law at all...” (6:35). This reminds us of the reciprocity thesis. The thought of a causality in accordance with the moral law “flows from” the thought of a causality free from nature. Thus, it might be argued either that Kant contradicts himself even in the same work or that the difference between his arguments is not fully appreciated.

In any case, Kant’s constant identification of freedom with the moral law gave rise to metaphysical debates on free will during his era. The problem of free and immoral action originated from these debates. In the next section, I summarize Johann August Heinrich Ulrich’s and Carl Christian Erhard Schmid’s reception of

⁹ *Kant: Critique of Practical Reason*, trans. by Mary Gregor, Cambridge: Cambridge University Press, 2015 (2nd ed.).

Kant's views between the years 1788 and 1790. Then, I present Reinhold's response to Ulrich and Schmid.

2.2 Ulrich, Schmid, and Reinhold

The imputability problem was first articulated by Ulrich and then by his student Schmid. Both were deeply influenced by Kant's philosophy. In *Eleutheriology or On Freedom and Necessity* (2022a), Ulrich argued that a deterministic conception of the world and of the causality of will follows from Kant's philosophy. He understood determinism as the idea that every effect has a "decisive" reason (p. 23). By the idea of indeterminism, he understood "true arbitrariness" (p.13). In an indeterministic system, there are no laws that connect effects to their causes necessarily, and there are no decisive reasons that justify why the same cause gives rise to the same effect under the same conditions. Ulrich also argued that there must be not only natural necessity but also moral necessity. A rational agent has determining grounds for her action, which are not arbitrary but "necessitated" by reason. The will of a rational agent is determined by the law of reason. Ulrich understood the necessitation of the will by reason as a causal relation, and he claimed that when we do what is morally right, we could not have done it otherwise. Therefore, although we are free from natural necessity, we are subject to moral necessity.

The last point also led Ulrich to question how immoral actions can be determined by reason if the law of reason is the moral law. His proposed solution is to understand immoral action as resulting from a "lack of exercise" or "weakness" of reason (p.13). However, he is not clear about what makes the exercise of reason defective, given that he understood the relation between reason and will as a causal

relation. If the only way reason can determine the will is by the moral law, it seems that in the case of immoral actions, there is no determination by reason at all.

The discontent with Ulrich's solution to the question of immoral action led his student Schmid (2022b) to come up with a more radical view, also known as intelligible fatalism. Schmid endorsed Ulrich's idea that the will is always causally determined by reason (p.77). However, Schmid was more radical because he denied Ulrich's idea of immoral action as a defective exercise of reason. Accordingly, in the case of immoral action, there is no determination by reason. Immoral action is not free because it does not embody reason's "self-activity", and action expresses this activity only when one obeys the moral law (p.75).

Schmid also proposed the idea of intelligible obstacles (*Hindernis*) (p.76). Intelligible obstacles, for Schmid, are not found in nature but in the "intelligible world", that is, in reason's domain. The self-activity of reason is impeded by these obstacles in the case of immoral action. He holds that we cannot know what these obstacles are because they are beyond our cognition. While reason is not active in the case of immoral actions, it is still able to deal with these obstacles, and therefore reject being determined by obstacles foreign to its own self-activity. Therefore, while immoral action is not determined by reason in any case, reason is still responsible in the sense that it can always assert its activity against intelligible obstacles. However, it is still not clear why the fact that reason does not assert its activity in the case of immoral action does not still count as the self-activity of reason.

Karl Leonhard Reinhold was influential in the debate on Kant's conception of free will. He was a figure sympathetic to Kant's philosophy. In *Letters on the Kantian Philosophy* (1792), he wrote that there was a serious misunderstanding in the reception of Kant's view by the "friends of the Kantian philosophy," which were

Ulrich and Schmid (as cited in Guyer, 2024, p.155). For Reinhold, their all-encompassing determinism was a misunderstanding of Kant's theory. Accordingly, such a determinism posed a threat to the reasonableness of Kant's ethical theory. He considered Schmid's intelligible fatalism to sidestep the common sense idea that evil is a free act imputable to the wrongdoer. Reinhold's premise is that a will is impure when it is not determined by reason. If the law of reason is the causal law that determines the will, then free and immoral action is not possible since it would be the effect of an impure will. Reinhold believes that this conclusion follows if we explain the relation between will and reason as a causal relation as Ulrich and Schmid did:

From the confusion between the to be sure self-active but nothing less than free action of practical reason—which gives nothing but the law—with the action of the will, which acts as the pure [will] only insofar as it freely grasps this law—nothing less than the impossibility of freedom for all immoral actions must result. As soon as it is assumed that the freedom of pure volition consists merely in the self-activity of practical reason, then one must also concede that impure volition, which is not effected through practical reason, is by no means free. Some of the foremost writers from the Kantian school have actually attempted to prove against Kant himself that in the case of immoral actions not merely the occasioning but even the determining ground of volition are to be sought outside the person, and that the will is free only in moral actions. (ibid.)

Reinhold wrote that if freedom of the will consists in acting according to practical reason, which “gives nothing but the law”, then “impure volition”, or evil will, cannot be free. This conclusion follows from the determinist premise that the will is pure and free when its acts are “effected through practical reason” and its law. An impure will cannot be free because it does not act according to the law of reason. Reinhold's criticism concerns the deterministic premise that the relation between the acts of the will and the law-giving act of practical reason is a causal relation and that the relation must necessarily hold. The view shared by Ulrich and Schmid, for Reinhold, does not metaphysically allow the possibility of an evil yet free will, which can freely act against the law of reason.

In writing his review, Reinhold believed that Kant, in *Groundwork III*, offered “pure expositions”, and that this suggested the impossibility of free and immoral action (p.268). However, he considered it a fault to regard these expositions as giving the full definition of freedom. He accepted the idea that acting according to the moral law is one aspect of the freedom of will. However, he denied that freedom of the will consisted of only that aspect. For Reinhold, the relation between the will and practical reason is not causal but normative. Practical reason restricts the will by giving its law as a norm to be followed, and it is up to the will to decide whether it follows or deviates from the law. Therefore, freedom of the will, for Reinhold, consists in choosing for or against the moral law.

Reinhold’s goal was to defend Kant from the misappropriations of his theory. However, this defense took him one step further in articulating a point that Reinhold thought Kant left unclear. In *Groundwork III*, Kant argued from the negative concept of freedom to the positive concept of freedom, as the property of the will to be determined by a law of its own, which is the moral law. Reinhold instead wrote that the positive concept of freedom must rather be defined as the “self-determination” of the will through choice (*Willkür*) for or against the moral law (p.271).

Reinhold argued that unless we can regard the will as free to choose for or against the moral law, we cannot genuinely call it free. “Without the practical law,” or the moral law, as a norm on its conduct, the will would not be free because it would depend on desires given by nature (p.273). “Without the natural law of desires” and with only the moral law, the will would be “self-active”, but would not be a free will that can act on desires. So, Reinhold’s concept of free will, or what he called “absolute freedom” against Schmid, involved independence not only from the

law of nature but also from the moral law. Without this double-sided freedom, he concluded, the will could not be regarded as genuinely free.

2.3 Kant's response to Reinhold

Kant wrote *the Metaphysics of Morals* during a time when the debate between Reinhold and others was ongoing. Although Kant does not name to whom he is responding in the passage below, it is commonly accepted that Reinhold was his target.

But freedom of choice cannot be defined - as some have tried to define it - as the ability to make a choice for or against the law (*libertas indifferentiae*), even though choice as a phenomenon provides frequent examples of this in experience. For we know freedom (as it first becomes manifest to us through the moral law) only as a negative property in us, namely that of not being necessitated to act through any sensible determining grounds. But we cannot present theoretically freedom as a noumenon, that is, freedom regarded as the ability of the human being merely as an intelligence, and show how it can exercise constraint upon his sensible choice; we cannot therefore present freedom as a positive property. But we can indeed see that, although experience shows that the human being as a sensible being is able to choose in opposition to as well as in conformity with the law, his freedom as an intelligible being cannot be defined by this, since appearances cannot make any supersensible object (such as free choice) understandable. We can also see that freedom can never be located in a rational subject's being able to choose in opposition to his (lawgiving) reason, even though experience proves often enough that this happens (though we still cannot comprehend how this is possible). - For it is one thing to accept a proposition (on the basis of experience) and another thing to make it the expository principle (of the concept of free choice) and the universal feature for distinguishing it (from *arbitrio bruto* s. *servo*); for the first does not maintain that the feature belongs necessarily to the concept, but the second requires this. - Only freedom in relation to the internal lawgiving of reason is really an ability; the possibility of deviating from it is an inability. How can the former be defined by the latter? (6:226)

Kant responds to Reinhold's claim that we cannot understand a person as free unless she is capable of choosing against the moral law as well. As we have noted, Reinhold thinks that genuine freedom consists of the ability to choose between the moral law and the law of nature or self-love. Kant takes Reinhold's definition of freedom as based on the conception of freedom as the liberty of indifference (*libertas*

indifferentiae). In this conception, freedom is marked by indifference, or equal distance, to at least two options: the moral law and self-love. The capacity to choose between them is what defines our freedom.

In the passage, Kant does not reject the idea that we encounter examples of such a capacity in our daily lives. However, for Kant, this is a consideration of “the power of choice (*Willkür*) as phenomenon”, not the consideration of “free power of choice”, or free *Willkür*, as a “supersensible object”. Kant accepts that *Willkür* as phenomenon, or the human being as “phenomenon”, shows us this ability (*Vermögen*).¹⁰ He also adds that we cannot understand how such a capacity is possible. Kant contrasts free *Willkür* as a supersensible object with *Willkür* as phenomenon. The latter is what is observable about *Willkür* while free *Willkür* is not observable because it is a “supersensible object”. By supersensible object, Kant means an object that we cannot observe in the phenomenal world, but comprehend by reason alone.

Kant rejects the idea that a supersensible object can be defined (*erklärt*) by a sensible object. The German word *Erklärung* does not have a fixed translation for Kant’s usage. It is usually translated as “definition”, “explication”, “exposition”, “explanation” and “declaration”.¹¹ In the passage above, Kant argues that a concept can be defined only with features that necessarily belong to it. He responds to Reinhold that acting against the moral law is not a necessary feature of the concept of freedom because only experience shows us such a capacity. The necessary feature of the concept of freedom is to act in accordance with the moral law. Kant also construes this feature as the capacity of the human being to act “in relation to the

¹⁰ *Vermögen* is translated both as “ability” and “capacity”. I will use them interchangeably throughout this thesis depending on the cited translations of Kant’s works.

¹¹ Kant notes the variation in expression in the first *Critique* (A730/ B758).

internal lawgiving of reason”. Deviation from the law of reason is the absence of this capacity, or an incapacity (*Unvermögen*); it is not, for instance, a misuse or a defective exercise of it, as Ulrich proposed. That is why Kant rejects including the capacity to choose against the moral law in the concept of freedom.

Another important point in the passage is the idea that “we cannot theoretically present (*darstellen*)... freedom as a positive property”. Kant refers to the positive concept of freedom, which is the concept that specifies how a free being acts. The definition of free *Willkür* depends on the positive concept of freedom for Kant. A theoretical presentation requires intuition, which can be intellectual or sensible. Given that a sensible intuition cannot yield theoretical knowledge about freedom, only an intellectual intuition can do so. However, Kant argues we do not have such an intuition (A260/B294). He also writes in *Metaphysics of Morals* that the concept of freedom is “positive” from the “practical standpoint” (6:214). When we think about what to do in practical deliberation, the concept of freedom is not only negative but also positive or action-guiding. Kant believes that reason imposes the constraint of willing maxims as universal laws when we act. In this respect, the concept of freedom, and its necessary feature can be defined only in accordance with what reason tells us to do in every instance of action.

The idea that we have no theoretical resources to present freedom positively is significant in Kant’s reply to Reinhold in one respect. Reinhold argues that the positive concept should be understood as “self-determination” through choice (*Willkür*) for or against the moral law” (p.271). Kant’s response is that including *Willkür* in the concept of positive freedom would be a theoretical presentation of the positive concept of freedom. In that case, we would present “the capacity of the human being considered merely as an intelligence, and show how it is necessitating

with respect to his sensible power of choice (*Willkür*)”. In my understanding, the problem for Kant is that Reinhold’s definition amounts to an illegitimate explanation of the relation between the noumenal world and the phenomenal world. Neither observation nor any intuition that we possess grounds such an explanation.

Relatedly, it is noteworthy that Kant does not define positive freedom as the capacity to make a choice for the moral law over self-love. Elsewhere in the *Metaphysics of Morals*, for instance, he defines positive freedom as “the ability (*Vermögen*) of pure reason to be of itself practical”, and he continues, “But this is not possible except by the subjection of the maxim of every action to the condition of its qualifying as universal law.” (6:214). Kant does not believe that freedom can be defined as choosing morality over self-love either. Reason, inasmuch as it can be “practical of itself,” has only one ability, which is to provide the universal form of laws to maxims.

Reinhold was surprised to find out that his defence of Kant’s theory of freedom was not endorsed by Kant himself when Kant responded to him in the *Metaphysics of Morals*. In his review of the *Metaphysics of Morals* published in 1797, Reinhold wrote that he never intended to derive his own concept of free will from experience or appearances though he took experience to verify the idea of freely chosen evil (2022, p.248).¹² Accordingly, he intended to derive his definition of free will from the consciousness of both the moral law and the natural law of desire. As he did in his response to Ulrich and Schmid, Reinhold maintained against Kant the idea that without admitting the existence of these two laws as the determining ground of the will, the human being cannot be genuinely free.

¹² The English translation of Reinhold’s review is available in Noller (2022) between pp. 238-249.

CHAPTER 3

TWO POPULAR RESPONSES

The debate on Reinhold's objection has continued for more than two centuries and led to different views. Some accepted that Kant could not provide a satisfactory answer.¹³ Paul Guyer (2024), for instance, accepted the charge that there is a conflict between Kant's *Religion* and other works (p.160). Though Guyer believes that it is possible to provide an answer, he writes that Kant's response to Reinhold gives us "no good reason" to reject Reinhold's definition of freedom as the ability of the will to choose for or against the moral law. Lewis White Beck (1960) argued, for instance, that the objection was based on a confusion of Kant's terminology. Beck wrote that the problem of freely chosen evil disappears once we better understand Kant's conception of "volition" (p.205). In this chapter, my purpose is to present two of the most popular responses to Reinhold's objection in contemporary philosophy.

3.1 The capacity reading

One response to Reinhold is known as the capacity reading.¹⁴ Accordingly, we can distinguish between an act reading and a capacity reading of Kant's argument for the reciprocity between freedom and the moral law. The act reading, which has been the more classical interpretation, holds that freedom consists in acting from the moral law. In this reading, the reciprocity between freedom and the moral law is strict in the sense that while acting from the moral law counts as an exercise of freedom, deviation from the moral law does not. The capacity reading instead holds that

¹³ See Timmerman (2007) and Guyer (2024),

¹⁴ Bojanowski (2023), Paton (1948), Louden (2000), and Potter (1974) more or less argue for the capacity reading.

freedom consists in our capacity to act from the moral law. Whether we act or do not act from the moral law, on this reading, does not make us free or unfree. It is only our capacity for the exercise of autonomy, not the actual exercise of it, that makes us free.

This suggestion was first proposed by H.J Paton (1946), who comments that “apart from the merely linguistic question”, there is no reason to think that Kant meant more than a “mere recognition of the categorical imperative” as what makes us free (p. 213). Paton believes that Kant meant a rather loose, not a strict, reciprocity between the moral law and freedom is expressed by the principle ought-implies-can, which Kant endorses.

Kant refers to this principle ought-implies-can following his *Factum* argument in the second *Critique* that the moral law is a fact of reason (5:30). There, Kant denies that a deduction of the necessity of the moral law from freedom is possible because we cannot cognize freedom on its own. We can only cognize it through the moral law as the fact of reason (*Factum der Vernunft*). In his example of a person threatened by his prince with execution, if he does not give false testimony, Kant writes that no one can predict the actual choice of the man (Ibid.). However, for Kant, the person knows that he can refuse to give false testimony through the judgment one ought not to give false testimony. Through the consciousness of “ought”, he “cognizes freedom within” and judges that he “can” (Ibid.).

However, the modal “can” signifies only the capacity to overcome sensible impulses and conform to the moral law for its own sake. If we could say, like Paton implied, that ‘can’ signifies only the capacity to act according to the moral law, not the capacity to conform to it, then this would undermine Kant’s argument in that section. The reason is that the moral law “leads us to the concept of freedom” by

showing us that it is possible for us to “set aside” all “sensible impulses”, like desires, and act only from moral considerations (5:30). The only capacity that assures us of our freedom for Kant is the capacity to act from the law. Therefore, the principle ought-implies-can does not show that we our freedom can be understood as acting against the moral law as well.

Another capacity reading has been more recently proposed by Jochen Bojanowski (2023). For Bojanowski, the word “incapacity” or inability (*Unvermögen*) in Kant’s response to Reinhold must not be understood as the impossibility of a capacity to deviate from the moral law (p.41). As we have noted, Kant responds to Reinhold that “only freedom in relation to the internal lawgiving of reason is an ability, the possibility of deviating from it is an inability.” (MM 6:226).

For Bojanowski, we must regard the word incapacity as “deficient exercise ” of the same capacity, not its absence altogether (p.42). Noting that there is no clarity on Kant’s use of the word incapacity and its meaning, Bojanowski points out Kant already uses the word “possibility” in the phrase “possibility of deviating from it [the internal lawgiving of reason] is an incapacity” in the same sentence. Thus, Bojanowski suggests that incapacity does not need to imply impossibility in Kant’s usage.

Bojanowski further suggests that an analogy can be drawn between acting immorally as “incapacity” and a heart that cannot properly pump blood. When a heart fails to pump properly, it is still a heart by virtue of possessing the same capacity. Likewise, one can still possess a capacity without being fully successful at exercising it. “Volition” that deviates from the moral law fails but does not cease to be a volition (p.42).

It is highly possible that Kant means the absence of capacity by incapacity. The sort of incapacity that he mentions in the relevant passage is one that belongs to the human being as an “intelligible being”, not as a “sensible being” or phenomenon. The distinction itself suggests that Kant considered immoral action as a capacity only in relation to the human being as a sensible being or phenomenon, while as an incapacity in relation to the human being as a thinkable or an intelligible being. That is, there are two separate capacities, and the capacity to act according to the moral law belongs to the human being phenomenon. Given that the organ analogy applies to one entity or one capacity, it cannot be coherently made when we consider Kant’s words in his response.

3.2 The *Wille-Willkür* reading

Another popular response to Reinhold’s objection is also known as the *Wille-Willkür* response.¹⁵ According to this reading, Kant recognized the need for a clear distinction between *Wille* (will) and *Willkür* (power of choice) due to the debate between Reinhold and his contemporaries on the imputability of evil. While the same terminological difference is found across Kant’s writings, the *Metaphysics of Morals* contains Kant’s own distinction before the paragraph where he responds to Reinhold:

Laws proceed from the will, maxims from choice. In man the latter is a free choice; the will, which is directed to nothing beyond the law itself, cannot be called either free or unfree, since it is not directed to actions but immediately to giving laws for the maxims of actions (and is, therefore, practical reason itself). Hence the will directs with absolute necessity and is itself subject to no necessitation. Only choice can therefore be called free. (6:226)

¹⁵ This view is defended by Allison (1990), Beck (1960), Silber (1960)

Kant writes that the will is not “directed to actions”, but to “giving laws for the maxims of actions”. For this reason, it cannot be “free or unfree”.¹⁶ Maxims “proceed” from choice, which can be free or unfree because it relates to actions. The will, which Kant equates with practical reason, is “subject to no necessitation” as the law-giver, while choice is subject to necessitation by the will. While *Wille* gives the law, and therefore has a legislative function, *Willkür* executes the law through maxims. The response holds that *Willkür* is free to execute or not execute the law given by *Wille*.

Lewis W. Beck (1963) writes that the failure to distinguish *Willkür* from *Wille* in the context of Reinhold’s debate “has embarrassed many defenders of Kant, who have accepted the conclusion that a good will is a free will.” (p.203). For Beck, if we suppose that there is only one “function” of will and one kind of “freedom”, then we are led to the problem of freely chosen evil (p.205). Contrary to Reinhold’s criticism, in Beck’s account, it is not Kant’s equation of practical reason and will that gives rise to the problem. The problem is rather the consequence of the failure to see that the same faculty, whether we call it will or practical reason, has two functions, each expressing a different concept of freedom.

Beck argues that *Wille* and *Willkür* express the legislative and executive aspects of practical reason respectively. Freedom of *Wille* is autonomy, and it is responsible for giving laws for the maxims of choice. While Kant does not call *Wille* free or unfree, Beck calls *Wille* free or autonomous. Freedom of *Willkür* is spontaneity or negative freedom (p.198). As we have discussed in the previous chapter, the negative concept of freedom is the conception of freedom by what it is not. A negatively free will is one that is not determined by incentives given by nature

¹⁶ Kant uses “free will” (*freyen Willens*) in countless places. I discuss this point in the last chapter.

like desires. Beck supports this reading with Kant's account from *Religion* that incentives "coexist with the absolute spontaneity of the power of choice (*Willkür*)" only by being incorporated into maxims (6:24). Kant argues that incentives, on their own, do not motivate us into action or serve as reasons for action. For incentives to be reasons for us, we need to incorporate them into maxims in the form of "I will do Y to get X". Incorporation of incentives into maxims is the function of *Willkür* or the power of choice for Kant. Kant calls *Willkür* spontaneous because it is free from the influence of incentives while incorporating them into maxims.

In Beck's reading, while *Wille* presents the law to *Willkür*, the latter can obey or disobey the law through maxims by virtue of its spontaneity. The power of choice retains its spontaneity when it disobeys the law; but when it obeys the law "it gains in freedom", by being "autonomous" as well as "spontaneous" (1963, p.199). Beck's reading hinges upon the idea that both *Wille* and *Willkür* are interrelated aspects of the same faculty. For him, this explains why the law legislated by *Wille* is not foreign to *Willkür*. *Wille* binds *Willkür* because the law represents what the latter would obey if it were an infinitely rational or a "holy will" (p.199). However, since *Willkür* is not in such a perfect condition, the law of *Wille* is a constraint on *Willkür*.

According to Nelson J. Potter (1974), Beck's response does not go beyond giving a new formulation to the old problem of two concepts of freedom raised by Sidgwick (p.593). Potter argues that Beck's distinction between *Wille* and *Willkür* corresponds to the two concepts of freedom which are, as Sidgwick (1888) maintained, found across Kant's ethical writings. One is the more neutral conception of freedom that is applicable to both moral and immoral action, while the other is the one applicable only to moral actions. Potter believes that the *Wille-Willkür* distinction only renames these two concepts of freedom without explaining how one

can be free in one sense while unfree in another sense. Relatedly, Courtney D. Fugate (2012) also notes that how *Willkür* can be bound by the law that it is free to disobey remains unanswered by Beck's reading (p. 24). The most Beck's reading does, for Fugate, is to frame the relation between *Willkür* and *Wille* in a perfectionist way where freedom of the former can either be merely spontaneous by violating the law or become both spontaneous and autonomous by following it.

Another version of the *Wille-Willkür* response is given by Henry E. Allison. Allison (1990) contends that Kant does not present a consistent account of *Wille* and *Willkür* throughout his works (p.132). His reading is therefore more of a reconstruction that he thinks more plausibly follows from Kant. Like Beck, Allison maintains that *Willkür* must be understood as spontaneous or negatively free. However, according to Allison, Kant's writings allow that *Wille* can be understood more broadly as the will, the whole faculty of volition including the legislative and executive aspects. Alternatively, for Allison, *Wille* can be understood more narrowly as the legislative aspect of the will, which is the one that Kant talks about in the *Metaphysics of Morals*.

In more recent work (2020), Allison calls the first and broad sense *Wille1* while the second and narrow sense *Wille2* (p.452). Against Beck, Allison argues that only *Wille1* is free in the sense of autonomy and that *Wille2* is neither free nor unfree because it is law-giving. *Wille1* is autonomous because the will can be a law to itself only when both the legislative and executive aspects are considered. When the legislative aspect, or *Wille2*, gives the law, and *Willkür* follows it, *Wille1* becomes a law to itself (p.455). In the case of evil, *Wille2* gives the law to *Willkür*, and the latter, being spontaneous, does not follow it and therefore renders *Wille1* heteronomous. Compared to Beck's, Allison's reading better coheres with Kant's

official definition in which *Wille*₂ is neither free nor unfree whereas Beck takes *Wille*, without distinguishing, to be autonomous. While Beck holds that *Willkür* can either remain spontaneous or become both spontaneous and autonomous, Allison reserves autonomy to the will or *Wille*₁. Contrary to Beck, for Allison, whether the will becomes heteronomous or autonomous is equally the result of the spontaneity of *Willkür*.

As Allison himself notes, his interpretation is not in accord with Kant's doctrine in the *Metaphysics of Morals* (p.456). In the paragraph that follows his distinction between *Wille* and *Willkür*, Kant responds to Reinhold that "freedom of *Willkür*" cannot be defined (*erklärt*) as the "capacity" (*Vermögen*) to act against "the inner lawgiving of reason" (MM 6:227). For Kant, this definition collapses into what he calls the "liberty of indifference", which he understands as "the ability to make a choice for or against the law" (6:226). Such an ability is only suggested by experience, or by *Willkür* "as phenomenon", not free *Willkür*.

In my understanding as well, the spontaneity of *Willkür*, as given in *Religion*, does not relate to the definition of free *Willkür*. As a negative concept, spontaneity does not tell us how a free agent acts, but only how she does not act. *Willkür* is spontaneous because it incorporates desires into maxims without being determined by these desires. However, the principle or law by which *Willkür* incorporates incentives into maxims is not revealed by the concept of spontaneity. Free *Willkür*, for Kant, is not only spontaneous but also positively free because its maxims are determined in accordance with a law. Since this law is the moral law, *Willkür* can be defined as free when its maxims are not only spontaneous but are also determined by the moral law. Therefore, when Kant says that *Willkür* is spontaneous in *Religion*, he does not mean that this defines free *Willkür*. As I will argue in the fifth chapter, it is

only *Willkür* as phenomenon, not free *Willkür*, that can be spontaneous and evil for Kant. I will also argue that the human will is always autonomous or good and never heteronomous or evil for Kant.



CHAPTER 4

KORSGAARD'S RESPONSE: THE UNINTELLIGIBILITY OF EVIL

In this chapter, I examine Korsgaard's response to the version of the objection formulated by Henry Sidgwick. However, most of her points apply to Reinhold's objection as well. Korsgaard argues that Kant's views on the reciprocity between the moral law and freedom support the idea that evil is unintelligible, not that it is impossible.

For Korsgaard, Kant's argument for the reciprocity between freedom and the moral law is made from the standpoint of a being free from all inclinations. Why a completely free being would "abandon" her freedom for an "object of inclination" is unintelligible (1996, p.15). Yet, Korsgaard also argues that the standpoint of freedom is practical, not theoretical. While practically evil is unintelligible, we cannot theoretically conclude that it is impossible.

In the first section, I briefly explain the distinction between noumenon and phenomenon in Kant's philosophy. In the second and the third sections, I present Korsgaard's response on the basis of her readings of *Groundwork III* and the second *Critique* respectively. Finally, I compare Kant's accounts of the first *Critique* and the second *Critique* on the concept of the unconditioned and argue that it is possible to infer from Kant the idea that evil is practically unintelligible.

4.1 Noumenon and phenomenon

For Kant, intuition is required for the conscious representation of objects (B160). In Kant's system, time and space are not properties of things as they exist independently of our intuition. They are formal features of our intuition that structure

the world. The ideality of time and space makes possible a distinction: things as they appear to us, or phenomena, and things as they are in themselves, or noumena. A phenomenon is an object that we can intuit and represent as located in time and space. A noumenon is an object as it exists independently of time and space conditions. The word noumenon originally comes from the Greek word *noos*, which means mind. Noumenon means the object comprehended by the mind, not the senses.

Kant believes that knowledge is not possible without conscious and objective representation of objects. Objects of experience must be first given to our intuition. Due to the forms of our intuition, which are time and space, representation of objects is possible only spatiotemporally. We can represent and know objects only as they appear to us in time and space. However, noumena cannot be given to our intuition since they are conceptually objects that exist independently of time and space conditions. For this reason, Kant does not treat the concept of noumenon the way it was treated before him. He argues that noumena cannot be comprehended because the concept of noumenon is “problematic” (B343). It is problematic in the sense that we cannot prove or reject by using our intuition that objects may exist independently of time and space. The possibility that objects may have non-spatiotemporal existence remains when these objects are not regarded as the objects of our intuition.

For Kant, the problematic status of the concept of noumenon prevents its comprehension by our reason. Noumenon is merely thinkable and possible without contradiction when the ideality of time and space is accepted. In this sense, noumenon represents what cannot be comprehended or known by us. That is why Kant also argues that the concept of noumenon has a “negative” function (B343). By representing what is “unknown”, noumenon “designates the boundaries of our sensible cognition” (B345). In other words, it shows us the boundaries beyond which

our intuition cannot represent any object and therefore leads to knowledge. In this respect, Kant treats noumenon in the first *Critique* as thinkable only in the negative sense. It is thinkable as a possible object that is not the object of intuition.

Another essential feature in the distinction between noumenon and phenomenon concerns the concept of causality. In Second Analogy, Kant argues that “All alterations occur in accordance with the law of the connection of cause and effect.” (B233). Kant argues perception alone does not establish “the objective relation” among events. These relations are regulated by “the concept of relation of cause and effect”, as a “pure concept of understanding”. (B234). In *Prolegomena*, Kant writes that “even the pure concepts of the understanding have no significance at all if they depart from objects of experience.” (4:313).¹⁷ Kant rejects the possibility of establishing causal relations for objects that we cannot experience or noumena. The relation of cause and effect, and the application of the concept of causality, is only possible for objects that are given to intuition.

Kant also argues that reason always seeks the condition of every conditioned event in time and space. Yet, this search can go infinitely back due to the principle that every event is conditioned by a preceding event (A306/B363). In this case, the unity required for the knowledge of the world cannot be established because reason demands that there be an unconditioned for every series of conditioned events. For Kant, reason’s solution is to posit an unconditioned that is itself not conditioned by any event in time and space. An unconditioned that conditions events in the phenomenal world can only be found in the noumenal world because conceptually, only noumena can be unconditioned by events in time and space.

¹⁷ *Prolegomena to Any Future Metaphysics: That Will Be Able to Come Forward as Science: With Selections from the Critique of Pure Reason*. (G. Hatfield, Ed.) (2nd ed.). Cambridge: Cambridge University Press.

While the thought of an unconditioned in the noumenal world is necessary, it is not possible to apply the category of causality to noumena due to the limitations of our intuition. The application of the concept of causality to objects of intuition allows us to express relations in terms of laws. In the first *Critique*, Kant argues that the concept of causality cannot be applied to objects as they exist apart from intuition (B21). For this reason, it is not possible to go beyond a grounding relation between the two worlds in causal terms. Alternatively, as we will further discuss in the chapter, a supersensible order beyond the phenomenal world cannot be given. The grounding relation is a mere thought that reason problematically admits in order to put an end to the regress of events.

4.2 Korsgaard's *Groundwork* reading

Korsgaard (1996) argues that the source of the problem of free and evil will is an ontological reading of Kant's philosophy. According to her characterization, the reading assumes that we exist both in the noumenal world and the phenomenal world. Kant's views are supposed to show that the moral law is the law of our noumenal selves (p.1). Korsgaard finds this reading problematic and objects that it goes beyond the boundaries of what we can know and that it serves as the source of the objection against Kant's views. If we are noumena and our will is free to the extent we act according to the moral law, it follows that we are not noumena, and we are not free when we act immorally. Therefore, we can't be both free and evil. For Korsgaard, the problem comes from a failure to appreciate the difference between the two mutually exclusive employments of reason in Kant's view. She argues that once we can distinguish between these employments, the problem of evil

yet free will “disappears” (p.2). In what follows, I briefly present her distinction between the two employments of reason in relation to her Kant reading.

In Korsgaard’s Kant reading, the theoretical employment of reason concerns what can be observed and explained. She argues that we can know and explain things as phenomena and ourselves as phenomenal beings, such as biological, psychological, and social beings (p.28). For Korsgaard, the idea of freedom cannot be theoretically employed because freedom is not observable in the phenomenal world. In any instance of action, there is no way in which we can observe and know whether a person acts freely or whether she has free will. Thus, the concept of freedom cannot be employed in our explanations about ourselves and the world.

Korsgaard argues that the practical employment of reason does not concern explanation but “deliberation” and principles of action (p.15). She writes that “the deliberating agent, employing reason practically, views the world as it were from a noumenal standpoint”. By the “noumenal standpoint”, she does not mean that we are noumena as a matter of fact. She means that we act as if we are noumena and free from all empirical conditions. Practical deliberation for Korsgaard is impossible without self-chosen principles and the conception of ourselves as noumena or as free agents (Ibid.). Though we cannot prove whether we are free or not, the necessity of acting as a “deliberating agent” justifies a conception of the world in which we are free creators.

Korsgaard also argues that the practical and the theoretical employments of reason do not result in contradiction in reasoning because each employment has its own “domain” and cannot step into the “territory” of the other (Ibid.). Explanations about the world and ourselves as phenomenal beings made by the theoretical employment of reason have no practical or normative status. In the same way, the

concept of freedom and the principles involved in the practical employment of reason do not have explanatory status.

Korsgaard's reading of Kant's argument in *Groundwork III* applies the distinction between the practical and the theoretical employments of reason. She argues that Kant's exposition of the law of free will must be read along with his following remarks on the necessity of acting "under the idea of freedom". If we read Kant in that way, for Korsgaard, we are led to the conclusion that the moral law is the principle of free will from "the standpoint of pure practical reason".

In *Groundwork III*, after Kant concludes that the law of free will is the moral law, he also argues that we are bound by the moral law insofar as we think of ourselves as free. His point is that although we cannot theoretically know whether we are free, we "act under the idea of freedom" in a practical respect (4:447). We necessarily think and act "as if" we are free from competing desires in a practical sense. He further argues that insofar as freedom is a necessary thought for us, we are bound by the laws of freedom. The important thing to remember is that Kant has just shown that the law of free will is the moral law in the previous section. He now concludes that we are bound by the moral law insofar as we think of ourselves as free.

For Korsgaard, acting under the idea of freedom means that "we must regard our decisions as springing ultimately from principles that we have chosen, and justifiable by those principles." (p.5). When we act under the idea of freedom, we treat our reasons and decisions as "springing ultimately from our principles". In turn, we treat these principles as the principles "we have chosen". Korsgaard argues that reasons can always be explained by more "fundamental" reasons and principles. If I say, "I want to go to the park now," you can ask my reason for going there, and I can

respond, “I want to meet a friend there.”. But you can continue asking “Why do you want to meet your friend?”. Explanation in terms of reasons can go infinitely. There must be one ultimate principle that we choose. Still, it must be a principle or law about which it is “incoherent to ask why a free person would have chosen it”. Korsgaard writes that *Groundwork III* is an attempt to “show that the categorical imperative has this status.” (p.6).

Korsgaard grants Kant’s assumption that a will that is negatively free must act according to a law. She asks, “If it [the will] is free to make its own law, why can't it make any law whatever?” (p.5). In order to answer this question, she sets up a scene where the will must adopt a law and “commence operations”:

At the standpoint of spontaneity, the will must, in order so to speak to commence operations, choose a principle or a law for itself. Nothing provides any content for that law. All that it has to be is a law (p.8)

Korsgaard argues that the law of free will cannot prescribe any particular ends, such as happiness and wealth. In such a case, it is still possible to ask for the reasons for adapting those ends. Consequently, the explanation of reasons would continue to go back. For Korsgaard, if the self-chosen law of free will prescribes only the requirement to be a law, independently of the content of the law, then the regress of reasons stops (ibid.).

Korsgaard argues that the formal requirement to be a law is expressed in Kant’s argument by the Formula of Universal Law:

Suppose that it chooses the categorical imperative, as represented by the Formula of Universal Law. This formula merely tells us to choose a law. Its only constraint on our choice is that it have the form of a law. Nothing provides any content for that law. All that it has to be is a law. By making the Formula of Universal Law its principle, the free will retains the position of spontaneity. Or, to put it a better way, the argument shows that the free will need do nothing to make the Formula of Universal Law its principle: it is already its principle. The categorical imperative is thus shown to be the law of spontaneity. In a sense, the Formula of Universal Law simply describes the function or task of an autonomous will. The moral law does not impose a

constraint on the will; it merely says what it has to do in order to be an autonomous will at all. It has to choose a law. (1996, p.8)

For Korsgaard, the will does not justify its actions by appealing to inclinations and the particular ends suggested by these inclinations when it makes the Formula of Universal Law its principle. In this respect, the universal form of law guarantees that the actions of the agent are never justifiable by particular ends. So, Korsgaard's reading is essentially based on the idea that if it has to be a law, it can be a universal law. The other option is a law that prescribes by virtue of its content, but this means that the spontaneity of the will is threatened.

Rather than establishing the principle of free will, Korsgaard's reading reformulates Kant's view in different terms. The idea that "nothing provides content for that law" is present in Kant's reciprocity thesis in the second *Critique* as well. Kant writes that a "practical law" has its determining ground only in its "lawgiving form" (5:29). He writes that the "material" of a practical law will always be empirical objects suggested by desires. Assuming that the will is free from all "empirical conditions", he concludes that only the "lawgiving form" remains as the determining ground of free will. In Korsgaard's terms, all that the will has to do in order to remain free is to give a law.

For Korsgaard, Kant's view *Groundwork III* must be understood in relation to the domain of practical reason. Insofar as we think of ourselves as free from the influence of incentives, just like the free will in Kant's account, reason leads us to the formal requirement that our maxims have a universal form, whatever incentive we choose to act on. In this practical respect, all reasons are universally shareable. Korsgaard holds that evil is therefore unintelligible or unthinkable from "the standpoint of pure practical reason". The reason why someone acts from maxims that do not have a universal form cannot be explained from the standpoint in which the

thought of freedom leads us to the moral law. Alternatively, in Korsgaard's words, why a free being would "sacrifice her freedom for some mere object of inclination is completely unintelligible." (p.15)

Korsgaard notes that Kant's argument is based on the idea of a pure will completely free from empirical conditions. However, we do not have "pure will" from the theoretical standpoint (ibid.). While it is not possible to give reasons for immoral maxims from the practical standpoint, we can explain them theoretically. For instance, we can observe that people cannot always resist their desires. To show that there is no real contradiction, Korsgaard appeals to the mutually exclusive functions of practical and theoretical employments of reason. As we have noted, in Korsgaard's Kant reading, theoretical explanation and practical justification cannot mix. Korsgaard argues that we cannot legitimately infer that evil is theoretically impossible from the idea that it is unintelligible from the practical standpoint (ibid.).

Korsgaard writes, "From the explanatory standpoint of theoretical reason, nothing is easier to understand than that a human being might evade duty when it is in conflict with her heart's desire.". While evil is unintelligible in the sense that it is not possible to see why a free will would choose it, this does not conflict with the explanation that people sometimes choose desires over duty and act on evil maxims. Yet, Korsgaard also argues that apart from what we can observe, the possibility of evil will remains on the noumenal level as well by a discussion of Kant's second *Critique*. In the next section, I briefly explain Kant's account of practical postulates, and then I continue with presenting Korsgaard's response.

4.3 Freedom as a postulate

Kant defines a postulate of practical reason as a necessary and practical belief for us to obey the moral law. God, immortality, and freedom are the three postulates that Kant discusses in the Dialectic chapter of the *Critique of Practical Reason* (5:133). The first two postulates concern the conditions necessary for the achievement of the object of the moral law, which Kant calls the highest good. The postulate of freedom does not concern the object of the moral law, but its ground in the sense that without believing that we are free, the moral law cannot apply to us. While the content of the postulate of freedom is theoretical in nature, that we are free, reasons for believing in our freedom are practical. For Kant, without the need to prove or demonstrate our freedom, we believe in our freedom for the sake of the moral law. Without this belief, we cannot obey the moral law.

These postulates are those of immortality, of freedom considered positively (as the causality of a being insofar as it belongs to the intelligible world), and of the existence of God. The first flows from the practically necessary condition of a duration befitting the complete fulfillment of the moral law; the second from the necessary presupposition of independence from the sensible world and of the capacity to determine one's will by the law of an intelligible world, that is, the law of freedom; the third from the necessity of the condition for such an intelligible world to be the highest good, through the presupposition of the highest independent good, that is, of the existence of God. (Ibid.)

Kant's remark above defines the postulate of freedom "as the causality of a being as it being insofar as it belongs to the intelligible world.". The postulate of freedom concerns noumenal or intelligible causality. Obedience to the moral law requires "the presupposition of independence from the sensible world". Without believing that we are free from natural causality, the moral law cannot apply to us.

Alternatively, without presupposing such independence, we cannot hold ourselves responsible for fulfilling the commands of the moral law. In that respect, we must

believe that we possess intelligible causality, which is uninfluenced by the laws of nature in the phenomenal world.

The second aspect of the postulate of freedom is “the capacity to determine one’s will by the law of an intelligible world, that is, the law of freedom” or the moral law. To obey the moral law, belief in independence from the laws of nature is not enough. We must also believe in our capacity to act from the moral law. This requires treating the moral law as “the law of an intelligible world, that is, the law of freedom”.

According to Korsgaard’s (1996) reading, the postulate identifies freedom, “understood as noumenal causality”, with the moral law (p.25). This identification is made on the practical ground that we must obey the moral law. When we regard ourselves as capable of obeying the moral law, we regard ourselves as noumena acting according to the moral law. However, this ground is not theoretical, namely that we are noumena, and as noumena, we act according to the moral law.

Freedom as a practical postulate does not reflect the idea of noumena acting according to any other law or principle than the moral law in Korsgaard’s Kant reading (Ibid.). This leads to the similar conclusion that evil is unintelligible in a practical sense. However, Korsgaard likewise argues that we cannot legitimately infer from a practical postulate the conclusion that an action cannot be free and immoral, or that there cannot be an evil free will in the noumenal world. There are two reasons for the illegitimacy of this inference in her view.

The first one is related to Korsgaard’s premise that the theoretical and the practical employments of reason cannot be mixed. The belief in the moral law as the law of the noumenal world is a “rational belief” necessary on the ground that we must obey the moral law (ibid). However, the belief does not present a ground for the

theoretical employment of the concept of freedom. It does not give us a ground for concluding that we are free, and as such we are governed by the moral law.

The second reason is that all theoretical explanations are restricted to the phenomenal world. However, the claim that a free and evil will is impossible is a theoretical claim about the noumenal world since freedom concerns the noumenal world. In Korsgaard's reading, Kant's arguments in *Groundwork III* and the second *Critique* are not based on a cognition of the noumenal world. Hence, they are not meant to prove that as noumena we conform to the moral law. Theoretically, we cannot know how the noumenal world and therefore free will can operate. For Korsgaard, the idea of a free and evil will must be theoretically left open though we cannot explain or prove it in any sense.

Korsgaard does not directly ground her reading on the distinction between the pairs noumenal-phenomenal and sensible-intelligible in Kant's terminology. However, her reading of these pairs in a different context also suggests the idea that evil is unintelligible while still possible on the noumenal level:

...I think it is significant that in the *Groundwork* Kant uses the language of "intelligible" and "sensible" rather than that of "noumenal" and "phenomenal"; and also that he changes his language in the *Critique of Practical Reason*. As I understand these terms, the noumenal world is the intelligible world insofar as it is thinkable. If we think of noumena at all, we must think of them as acting in the only way that is intelligible to us, which is according to the laws of freedom. But at the same time we must always admit the possibility that the noumenal world is unintelligible to us...The moral law characterizes noumena insofar as they are intelligences (insofar as we can think of them) because acting according to it is the only thing it makes sense for them to do, and this is already a normative point. (1992, p.330)

Korsgaard's reading involves a "conceptual" relation between the terms noumenal and intelligible.¹⁸ In Korsgaard's Kant reading, the noumenal world is not knowable,

¹⁸ Allison's (2020) suggestion on how to read the noumenal-intelligible is also similar (p.282). However, as I will show, Korsgaard's reading of the distinction goes beyond a conceptual relation between the pairs.

but it is thinkable, and to that extent, it is the intelligible world. She reads the intelligible world as the conception of the noumenal world in accordance with “the laws of freedom”. Her premise here is that the moral law is the law of freedom. When we “think of noumena,” we think of them as “intelligences” acting in accordance with the laws of reason. In Korsgaard's reading, intelligences are essentially noumena, but ones that we think of as in accordance with the laws of freedom. In other words, the idea of an intelligible world has a normative or practical import in Korsgaard's reading. The noumenal world represents an unknowable world, while the intelligible world represents what we can think of the noumenal world for practical purposes.

Korsgaard also writes in the quotation that “we must always admit the possibility that the noumenal world is unintelligible to us.”. It is possible that noumena may not conform to the laws of reason. While as intelligences they necessarily conform to the moral law, it is open that as noumena they may act differently. For this reason, it follows from Korsgaard's reading of the distinction that a free and evil will is still possible in the noumenal world though this circumstance is unintelligible practically.

In Kant's writings, it is not possible to map out a clear distinction between the terms noumenal and intelligible. In the second *Critique*, for instance, Kant usually uses the pair sensible-intelligible instead of the pair phenomenal-noumenal. Hence, it can be said that Korsgaard's reading of these pairs does not directly follow from Kant's usage of the terms. However, there is a way to make sense of the idea that evil is unintelligible in a practical sense by comparing Kant's accounts between the first *Critique* and the second *Critique* on “the concept of the unconditioned”.

Now after speculative reason has been denied all advance in this field of the supersensible, what still remains for us is to try whether there are not data in

reason's practical data for determining that transcendent rational concept of the unconditioned, in such a way as to reach beyond the boundaries of all possible experience, in accordance with the wishes of metaphysics, cognitions a priori that are possible, but only from a practical standpoint. (B21)

As we have noted, the idea of an unconditioned in the noumenal world is a necessary thought of reason. In the B edition of the first *Critique*, Kant reiterates that transcendental idealism allows us to think of the unconditioned without contradiction. Then, he argues that we cannot go further than this mere thought and think of a supersensible order for the noumenal world in the way we can think of a sensible order for the phenomenal world, which is nature.

If theoretical reason could know what is beyond the sensible order of things, then the “transcendent rational concept of the unconditioned” could also be “determined”. In other words, we could cognize the unconditioned and its effects in the phenomenal world. While Kant denies that knowledge of this is possible within the limits of theoretical reason, he hints at the second *Critique*, the task of which is to argue that the concept of the unconditioned can be determined “from a practical standpoint.”. In this respect, one of Kant’s goals in the second *Critique* is to argue that *pure* practical reason determines the concept of the unconditioned with the idea of causality through the moral law:

I grant the mechanism of natural necessity the justice of going back from the conditioned to the condition ad infinitum, while on the other side I keep open for speculative reason the place which for it is vacant, namely the intelligible, in order to transfer the unconditioned into it. But I could not realize this thought, that is, could not convert it into cognition of a being acting in this way, not even of its mere possibility. Pure practical reason now fills this vacant place with a determinate law of causality in an intelligible world (with freedom), namely the moral law. (5:49)

Kant argues that the intelligible cause, which lies behind all the appearances in the world, cannot be theoretically cognized. While the first *Critique* leaves open the idea of an unconditioned, the concept of causality cannot be applied to it. Therefore, the

mere idea of the unconditioned does not give us a first cause whose effects are encountered among appearances. Theoretical cognition is restricted to events that operate under “natural necessity”. That is why Kant also writes in the same place that the idea of an intelligible world is theoretically “indeterminate” and “problematic” (Ibid.).

Kant argues above that only pure practical reason presents a ground for turning the idea of an unconditioned into the cognition of a free being acting in a certain way. Kant ties the possibility of this cognition to “a determinate law of causality”, which theoretical reason could not provide. Practical reason makes this cognition possible by providing the required law of causality, which is the moral law. In this case, the concept of causality is not used for theoretical purposes for the objects of senses, but for the practical purpose of determining the actions of the agent. Kant calls this “the supersensible use” of the concept of causality (KpV 5:6). Through the supersensible use of the concept of causality, the unconditioned is thought determinately in the actions of an agent whose effects are observable in the phenomenal world.

I believe that Kant’s account in the second *Critique* supports the idea that free and immoral action is unintelligible practically. Cognition of freedom requires a “determinate law of causality” in accordance with which a free being acts.¹⁹ In other words, our cognition of freedom for Kant is the cognition of the action of a being in accordance with the moral law.

¹⁹ By cognition, Kant refers to practical cognition, not theoretical cognition. Theoretical cognition is possible when the object of cognition is provided by senses. Practical cognition is possible with objects provided by pure practical reason, which are actions possible through freedom” (5:58). I will talk more on practical cognition in the next chapter.

Put differently, the intelligible world conforms to the moral law from the practical standpoint. Therefore, Korsgaard's reading that the intelligible world is the practical conception of the noumenal world seems to have a place in Kant's thinking. Though we cannot theoretically know how the unconditioned in the noumenal world operates, the unconditioned is thought lawfully in a moral agent acting in the phenomenal world.

In the next chapter, I offer an alternative reading of Kant's posthumously published notes. In his notes, Kant argues that the human will is good, and that we cannot explain evil "from the intelligible substrate of free will". I will argue that Kant's response implies that evil is unintelligible because why we would act in the wrong way can only be explained in terms of empirical conditions, and therefore that evil can be explained as a fact about the phenomenal world.

CHAPTER 5

THE GOOD HUMAN WILL

In this chapter, I offer a reading of Kant's Response in *Preliminary Notes and Reflections to the Introduction to the Metaphysics of Morals, Before 1797*.

Preliminary Notes is Kant's posthumously published two-page incomplete work consisting of notes and reflections on the topics of the *Metaphysics of Morals*. Kant's notes have not received sufficient attention in the context of Reinhold's objection. I aim to present a reading of Kant's notes in harmony with his official writings, especially his official response to Reinhold in the *Metaphysics of Morals*.

In his notes, Kant maintains his previous response that only experience shows us a capacity for choosing evil, and that freedom cannot be defined over this capacity. Kant argues that only *Willkür* can be evil, and "Wille is good in all human beings" (2022, 230). While Kant's point on *Willkür* is already present in his official answer to Reinhold in the *Metaphysics of Morals*, the idea that the human will is good is not usually encountered in his official writings. In this chapter, I aim to interpret this idea as a crucial aspect of Kant's response to Reinhold.

In making his objection, Reinhold assumed that the human will can either obey the moral law or deviate from it. I argue that for Kant, the human will is the source of the moral law in the sense that it never fails to give the universal form to the maxims of choice even when we act against the moral law. However, as I will discuss, Kant rejects Reinhold's implicit premise that the human will can be evil.

My reading will also try to answer whether there is a tension between the fact that we act immorally and that the human will is good by employing the language of worlds. In doing so, I will be arguing that Kant's implicit assumption in his response

is that the will is part of the noumenal world. In parallel with Kant's response that freedom cannot be defined by "appearances" (MM 6:226), I argue that our encounter with evil in the phenomenal world cannot explain how the will operates in the noumenal world.

My other purpose is to discuss Kant's view that why we act immorally "despite the good will" cannot be explained (*erklärt*) in terms of freedom, but "mechanically" (p.231). This unusual remark from Kant completes his response to Reinhold in one significant way. While his official response already stresses the idea that we encounter evil in the phenomenal world, he does not tell us how it can be explained, if not as freely chosen. I read Kant as arguing that since our acquaintance with evil is restricted to the empirical conditions, why we act morally can be explained only mechanically, that is, as a fact about the phenomenal world. Relatedly, I will also suggest that we can explain why people act immorally not in relation to their will, but in relation to empirical conditions that influence their choices.

This chapter also continues to discuss the idea that evil is unintelligible with Kant's remarks on evil actions being explainable only mechanically. Kant's responses imply that any explanation of why we act immorally cannot legitimately extend beyond the phenomenal world. I take this to suggest that we cannot explain evil actions in terms of reasons that arise from the will, but mechanically or in terms of observable conditions that influence us. I conclude that Kant's response to Reinhold is not merely an issue of definition or explanation. It is one that speaks in favor of the idea that why people act badly only finds an answer in the conditions of humans and the world.

5.1 Kant's notes and *Wille-Willkür*

Kant's first remark in the *Preliminary Notes* extends his distinction between *Wille* and *Willkür* in the *Metaphysics of Morals*. In the latter, Kant called *Wille* as the source of laws and identified it with practical reason. Now he likewise writes that *Wille* is "the lawgiver for the power of choice (*Willkür*) and is also " the absolute practical spontaneity in the determination of the power of choice." (2022, 230). However, he also adds in an interesting way, "Precisely for that reason it [*Wille*] is good in all human beings and there is no unlawful volition."

The idea that the human will is good, or human volition is always lawful, is not one that Kant usually makes. In his famous illustration of a "scoundrel" in *Groundwork*, who "transgresses" the moral law, Kant writes that the scoundrel has an "evil will (*bösen Willen*) as a member of the world of sense" (4:445). Yet the scoundrel is still "conscious of a good will" when he "believes" that he is also a "member of the world of understanding". Even in the moment of transgressing the moral law, he is "cognizant" of its "authority". So, Kant rejects an ultimately evil will that does not recognize the authority of the moral law. However, in his notes, Kant seems to deny evil will or unlawful volition in humans altogether. As I will further argue, the idea that the human will is good constitutes Kant's lesser known response to Reinhold's objection and especially his implicit premise that the will can be evil.

Following his remark on the goodness of human volition, Kant argues that evil is rather found in "the maxims of the power of choice":

But the maxims of the power of choice can be evil because they relate to actions as appearances in the sensible world, and the power of choice (as a natural faculty) is free in view of those laws (of the concept of duty) through which it is actually not immediately determinable, but rather only by means of the maxims it adopts in conformity with or contrary to the concept of duty. (2022, p.230)

Kant defines *Willkür* as “a natural faculty” whose maxims concern “actions as appearances in the sensible world.”. While the human will does not concern maxims, but laws for maxims, *Willkür* as a natural faculty can be observed in the sensible world, or as Kant also writes in the same text, in the human being as “phenomenon” (p.231). This much follows his previous response to Reinhold except the point that *Willkür* can be evil because its maxims are not “immediately determinable” by “duty” but by virtue of maxims that can be “in conformity with or contrary to the concept of duty” (ibid.,). Kant calls freedom in this sense “spontaneity”.²⁰

The power of choice is the capacity to choose between given objects. Thus, the opposition between these objects must concern a relation in accordance with laws of sensibility[.] Hence, this is itself already an evil power of choice. The ground of the possibility of a power of choice as such in the concept of the human being as noumenon is only that of freedom (independence from determinations by sensibility, [and is] therefore merely negative). (2022, 230-231)

Kant defines *Willkür* as “the capacity (*Vermögen*) to choose (*wählen*) between objects (*Gegenständen*)”. He argues that the competition between these objects, such as ones presented by inclinations, is made possible “in accordance with laws of sensibility”. The conflict between forces in human nature, along with the capacity of *Willkür* to make a choice between objects of inclinations, shows us that *Willkür* can be evil. In his remark, Kant writes that the ground of the possibility of an evil *Willkür* “in the concept of the human being as noumenon” is spontaneity or negative freedom.

In my understanding, the key point here is that the negative conception of noumenon, or freedom as spontaneity, does not disclose the essence of the concept of freedom or explain the noumenal world. What follows from Kant’s remark is that the

²⁰ In *Religion* as well, Kant calls *Willkür* spontaneous, however not in relation to immoral maxims, but in relation to its general capacity to “incorporate” incentives into a maxim without being determined by them.(6:24).

human being as noumenon, or freedom, is thinkable to us positively in the idea of a being acting according to the moral law in the phenomenal world. Thus, Kant's response to Reinhold is that the possibility of acting against the moral law cannot be included in the concept of freedom or "in the concept of the human being as noumenon". In familiar terms, insofar as we can think of freedom as the form of our actions, or equally think of ourselves as noumenon, what is thinkable is that we act according to the moral law. Relatedly, the judgment that wrongdoers are spontaneous regarding their particular choices in the sense that they are not determined by external causes, for Kant, is not accompanied by the thought of a lawful causality expressing their spontaneity. Given that freedom as a causality is always lawful, Kant is against calling a spontaneous and evil *Willkür* a free *Willkür*.²¹

In considering Kant's official response to Reinhold as well, we have two conceptions of *Willkür*, one is *Willkür* as a supersensible object and the other is *Willkür* as a natural faculty or phenomenon. The former always acts on maxims that comply with laws given by *Wille* while it is possible for *Willkür* to deviate from *Wille*. In either case, *Wille* or the human will is always good for Kant. Kant also writes that the independence of *Willkür* from duty cannot be "defined (*eklagt*) as the subjective possibility of deciding in conformity with or contrary to the law, [deciding on] the unlawfulness of actions whatsoever, for that would be tantamount to an evil will." (Ibid.).

²¹ "Freedom is thus not at all a faculty for choosing evil, but rather the good, because our reason commands only the good." in *Lectures on Metaphysics*, ed. by K. Ameriks and S. Naragon, trans. by K. Ameriks, Cambridge: Cambridge University Press, 1997, 29:903.

5.2 Goodness of human will and evil maxims

Kant accepts that we choose to act on immoral maxims, however, he rejects that we can understand the human will as the source of these maxims. Only an “evil will” would be able to give rise to these maxims and determine *Willkür* accordingly by “deciding on the unlawfulness of actions.” In my reading, what Kant rejects is not that we decide to act immorally, but that we act immorally for its own sake, which is what Kant elsewhere calls diabolical evil.²² We can act from the moral law, or for the sake of good, and in that case, nothing empirical such as interests and social factors, are motivating for us, except the morally good end itself. Yet, we cannot act for the sake of evil in the same way. As I will argue with further premises, the given considerations are among the assumptions behind the slogan of Kant’s notes that the will or *Wille* is good. After defining the freedom of will as a positive property in comparison to the mere spontaneity of *Willkür*, Kant argues that though experience shows us that we have evil maxims, the human will is still good.

But the will is free in another way because it is lawgiving and not obedient to the law of nature, nor any other law, and to this extent freedom is a positive capacity not to choose (for here there is no choice) but rather to determine the subject with respect to the sensible component of action. The basis of the circumstance that this capacity does not always result in the determination of the power of choice to the good, but rather evil actions and maxims arise despite the good will cannot, as phenomenon, be explained from the intelligible substrate of free will. (p.231)

One important idea that Kant reiterates is that only experience shows us “evil actions and maxims”, but he adds in an unusual manner that the will (*Wille*) is still good. Alternatively, the fact that we act immorally does not show that the human will is evil. Why *Willkür* is not always good, “despite the good will...cannot be explained from the intelligible substrate of free will.” As his previous sentence also reveals,

²² See *Religion*, 6:35

Kant's thinking follows the argument of *Groundwork*, according to which a free will and a will that follows the moral law are the same. Evil is unintelligible insofar as free will is concerned, because "from the intelligible substrate of free will", not only the human will, but also *Willkür* is good.²³

Kant is never clear about the difference between *Wille*, or the human will, and free will (*freien Willens*). If we consider his account in the *Metaphysics of Morals*, and also as indirectly suggested by *Preliminary Notes*, *Wille* is not free because it is the source of the moral law, so it is not "under" the moral law (MM 6:226). By this account only *Willkür* can be "free or unfree" (ibid.). Yet, he now argues that the will is free "because it is lawgiving and not obedient to the law of nature" (2022, p.231).

In both works, the human will is the source of laws for maxims and does not directly relate to actions, and it never fails to be lawgiving for *Willkür*. For this reason, it can be said that Kant calls *Wille* good will rather than free will in *Preliminary Notes*. Both the *Groundwork III* and *Preliminary Notes* define free will as the human will in relation to actions in the sensible or phenomenal world. In the latter work, freedom of will is a "positive capacity" that determines "the sensible component of action" in accordance with the moral law. Therefore, "from the intelligible substrate of free will", only good maxims can arise and likewise good actions can be explained. The other circumstance, in which "evil actions and maxims arise", cannot be explained in relation to free will.

²³ One rare instance where Kant mentions "intelligible substrate" is found in his discussion of the compability of freedom and deterministic explanations of actions (KpV 5:100). However, his discussion does not concern the intelligible substrate of free will, but of humans in general. The intelligible substrate serves as the conception of the human being as the cause of actions which are in accord with the natural law of causality in the phenomenal world. I believe that the intelligible substrate that Kant discusses in relation to free will has the same function. It regulates a conception of the human will (*Wille*) as the source of actions as appearances in the world.

The language of two-worlds presents a way for better understanding what Kant might mean by free and its intelligible substrate. Since the human will is the source of the moral law and does not directly relate to actions in the phenomenal world, it is better treated as part of the noumenal world. Secondly, it more directly follows from Kant's definitions that free will is the human will as an efficient causality in the phenomenal world. Then the intelligible substrate of free will concerns the "property" of the will to determine "the sensible component of action" or alternatively make a change in the phenomenal world.

If we take the human will as part of the noumenal world, then the intelligible substrate of free will can be said to represent our membership both in the noumenal and the phenomenal worlds. In this sense, the intelligible substrate is the bridge between the causal determination in the phenomenal world and freedom in the noumenal world. While the action is causally determined in the phenomenal world, it is at the same time freely performed by the same being as noumenon or as a member of the noumenal world. Kant's response seems to convey the idea that we can know or relatedly explain ourselves as members of both worlds insofar as we act according to the moral law.

The last point seems too strong, for if we are not also the members of the noumenal world as beings who can act against the moral law, it might be thought that we are causally determined in the phenomenal world when we act immorally. To remember, Kant argues that *Willkür* is spontaneous, that is, free from the influence of "sensibility". Yet, this is a judgment that we arrive at about ourselves and others when we "consider" ourselves as phenomenon.²⁴ For this reason, the spontaneity of

²⁴ What I particularly have in mind is Kant's remark that "The human being considers himself, as far as his power of choice is concerned, as phenomenon and to this extent is subject to laws regarding the form, i.e. the maxims, of his actions in which he has a choice. This freedom signifies nothing more than spontaneity. Hence the power of choice is free to do or refrain from doing what the law

Willkür does not by itself tell anything about freedom, or the noumenal world. While Kant takes immoral actions to be imputable to the spontaneous *Willkür*, the definition or explanation of freedom, or free *Willkür* as a “supersensible” or noumenal object, cannot be based on what we observe about *Willkür* as phenomenon. Relatedly, the issue is not whether we are causally determined or not when we act immorally in a metaphysical sense, but how we can know and explain such actions.

Kant’s idea in his official response to Reinhold that what we can observe as phenomenon cannot explain or define the concept of freedom is also relevant here. If we think in the language of two worlds, then it can be equally said that what happens in the phenomenal world cannot explain the noumenal world. This idea, I believe, is behind Kant’s contention that the fact that we sometimes have evil maxims, or act immorally, does not show that our will can be evil. The human will, or equally practical reason, never fails to give the moral law even when we act immorally. For this reason, the human will is good despite evil maxims. Then, what is left for the explanation of evil actions is an explanation that does not involve freedom or equally extend into the noumenal world. Following his remark on the intelligible substrate of free will, Kant writes that the explanation of actions “when they are evil” can only be made “mechanically”(mechanisch):

Just why we imagine what is outside us in space and what is inside us in time, and not the other way around, no reason can be given, because this concerns the sensual form of the objects, so it is here with the actions that we can only mechanically explain when they are evil, but never why such a mechanism is found in us. (p.230)²⁵

commands” (p.231). For Kant, when we “consider” ourselves as phenomena, we also judge that the power of choice is spontaneous in the face of the conflict between different objects, and not determined by the moral law as well.

²⁵ This sentence is missing in the English translation and present in original German text with the same page number, Kant, I. (1902–) *Gesammelte Schriften*, Akademie-Ausgabe, Band 23, *Vorarbeiten zur Metaphysik der Sitten*. Berlin: de Gruyter. I thank Jörg Noller, the translator of Kant’s *Preliminary Notes*, for kindly providing the translation of the sentence upon request.

This remark is not surprising because he has just rejected that free will can be the source of evil maxims. However, it is also surprising because he does not usually refer to mechanical explanations, particularly for evil actions in his other works. In the second *Critique*, Kant writes that as a “sensible being”, the actions of a human are “mechanically (*mechanisch*) conditioned” (5:101). Actions of the human in this respect are always the consequence of events in time and space (5:39; 5:97). Kant’s remark completes his previous response in one important sense. If our acquaintance with evil is restricted to the phenomenal world, or the human being as phenomenon, then the explanation of evil actions can only be mechanical. In other words, since the explanation cannot extend into the noumenal world, how and why we act immorally can be explained not in terms of volition or freedom, but in terms of causes that can be observed in the phenomenal world. Though Kant is not clear here, it is plausible to consider these causes as observable empirical conditions that relate to the world and the individual, such as biological, psychological, social facts and events of every sort in time and space. Each set of facts, such as the inclinations of the individual, can be explanatory depending on the type of questions asked about the agent and her actions.

Kant also notes that explanations of evil actions do not answer why the “mechanism” that leads us to evil “is found in us”. His analogy is with time and space as forms of intuition. We cannot explain why they constitute the shape of our inner and outer worlds in the way they do or serve as “the sensual form” of objects (2022, p.230). An answer is not possible without assuming a stance from which knowledge is not restricted to objects represented in time and space, which would be a noumenal point of view, or equally mean the presence of an intellectual intuition. In the same way, for Kant, our knowledge about evil is restricted to empirical

conditions of the human being as phenomenon. Any explanation of immoral actions falls short of telling us why we act immorally at all. This question can likewise be answered by assuming a stance from which our acquaintance with evil is not restricted to experience.

If my reading is correct, then the idea that evil is unintelligible in a practical or normative sense has a place in Kant's thinking. The fact that we can explain evil actions mechanically implies that evil is inexplicable in terms of reasons or "intentions" that are traceable to the freedom of the agent as their source.²⁶ Rather, these reasons are always traceable to empirical conditions. In this sense, reasons for evil are not genuine reasons at all, but at most explanations that refer to the mechanism that affects the individual in various ways. Moral actions for Kant, however, are not like that because they are chosen in accordance with a law that is not found in nature. For this reason, the moral law serves as the law through which reasons for action can be explained as totally belonging to us, not to nature. These points also remind us of Korsgaard's Kant reading. For Korsgaard (1996), we can explain evil as the result of a "conflict" between duty and inclinations where the human being "evades duty", but not as the result of the own principle of free will (p.15).

In the same way, Kant argues that immoral maxims admit no explanation in terms of freedom. He goes further than Korsgaard and removes any trace of reason or will in the explanation of immoral actions, which makes explanations completely

²⁶ Though Kant is rather talking about organisms and artefacts, I believe that his remark from the third *Critique* is also related point given that he takes evil actions as mechanically explicable: "Now if one asks why a thing exists, the answer is either that its existence and its generation have no relation at all to a cause acting according to intentions, and in that case one always understands its origin to be in the mechanism of nature; or there is some intentional ground of its existence (as a contingent natural being)" (5: 426). I thank Berna Kılınç (2011) for drawing my attention to Kant's remark by sharing with me her paper.

mechanical. It is not, for instance, the weakness or corruption of human will or reason, but empirical conditions that explain why we act immorally. While Korsgaard seems to be more agnostic about evil volition on the noumenal level, Kant's remarks deny such volition to humans.

In fact, the unfailing goodness of the human will is at the basis of the idea that evil is unintelligible or inconceivable. To see why, we must consider Kant's other alternative to mechanical explanation, which is an evil will that does not appeal to nature nor recognize the authority of the moral law at all. In effect, such a will would determine *Willkür* for the sake of evil not because of any empirical factors. Kant's premise here, I believe, is that no human wants to do evil and act for its own sake regardless of empirical conditions. So, evil is unintelligible in the sense that we cannot explain how and why we act immorally without appealing to empirical factors. Yet, why we do the morally right or good thing can be explained without the same appeal. Drawing the conclusion about the goodness of human will, and relatedly the unintelligibility of evil, depends on the premise that the human volition can be directed towards ends for their own sake. The following reconstruction puts that premise in place by employing the language of worlds.

1. As a part of the noumenal world, the will can determine *Willkür* in the phenomenal world without the influence of anything empirical, but for the sake of the end itself.
2. We know that we can act for the sake of good/we can act from the moral law alone.
3. We also know that we cannot act for the sake of evil.
4. Therefore, the human will is good.

The second and the third premises are Kant's obvious assumptions for which textual evidence is abundant. The first premise is what I take to be his more implicit assumption in his response to Reinhold. Kant's rejection of evil volition hinges on the idea that the human will, assuming provisionally that it can be either good or evil, can "decide on" the lawfulness or unlawfulness of actions. This is equivalent to the idea that the human will can be directed toward ends regardless of empirical factors and for the sake of these ends. That is why, in Kant's argument, the will is best treated as part of the noumenal world, or less metaphysically, as part of the noumenal view that we take of ourselves.

The second and the third premises concern not only the human will, but human agency more generally. Kant's argument takes it for granted that we can act for the sake of good ends because we can act from the moral law alone. His remarks against evil volition in humans reveal the other idea that we cannot act solely from evil motives, unlike the circumstance in which we can solely act from good motives. Since we can act for the sake of good, but not evil by virtue of our volition, the human will is good and cannot be considered as the source of evil.

Putting Kant's response in the language of worlds has the merit of explaining why evil is unintelligible. As we have noted, free will is what is thinkable or intelligible in the relation of the human will to actions in the world. A free will is not only the source of the moral law, but the source of the actions that are in accord with the moral law. In this respect, as beings with a free will, we are not only members of the noumenal world, but members of both worlds. It follows from Kant's discussions that good actions reflect our membership in both worlds because we can explain why we choose to act so regardless of empirical circumstances.

It also follows that insofar as we can explain how and why we act immorally, we are only members of the phenomenal world. Again, the issue is not whether we are causally determined or not. The issue is whether we can find the ultimate justification for acting badly in our volition. Kant's answer is no because a justification that does not appeal to our conditions and those of the world is possible only for morally good actions. No answer can be given why a free agent would act in the wrong way unless we are also willing to accept her as a member the phenomenal world.

Lastly, we can consider how Kant's separation between an evil *Willkür* and a good, or free *Willkür* relates to the unintelligibility of evil. Like Korsgaard, we can say having reasons for action requires freedom, which involves not only spontaneity but also a law from which a spontaneous being can derive her reasons (1996, p.5). Good *Willkür* is free because it justifies its reasons without appealing to the laws of nature, but to the moral law, or in Kant's words, simply because it follows the good human will. In other words, it expresses the standards arising from the will. Evil *Willkür* is free from the influence of laws of nature. Yet, such a *Willkür* has no law, or in Kant's words, a lawgiving "positive capacity" that determines the agent without appealing to the laws of nature (2022, p.231). In this sense, for Kant, when someone has chosen to act immorally, we judge that her choice is free. Yet, when it is a matter of volition and justification, the same judgment cannot be made from the *intelligible substrate of free will*.

CHAPTER 6

CONCLUSION

For Reinhold, the fundamental problem was that Kant failed to distinguish practical reason from the will. I have argued that, for Kant, the fundamental problem in Reinhold's thinking was that he failed to distinguish the human will from *Willkür*, or the power of choice. Kant argues that only *Willkür* can be evil. While this part of his answer has been commonly known, this thesis has attempted to draw attention to the less discussed part of his answer, namely that the human will is good.

Reinhold argued that freedom concerns the capacity to choose between good and evil because he assumed the human volition itself can be evil in the sense that it can reject standing under the moral law. I have argued that Kant's response points to the idea that an evil will is diabolically evil in the sense that it can give rise to immoral actions for their own sake, that is, regardless of empirical circumstances. Kant thinks that the human will never fails to give or be the source of the moral law for the maxims of choice. For this reason, it is good.

I have also argued that the distinction between the phenomenal world and the noumenal world is key to resolving the apparent tension in the coexistence of evil *Willkür* and good *Wille* in the same agent. For Kant, we know by observing ourselves and others that we make morally wrong choices in the phenomenal world. However, insofar as *only* our volition is concerned, we are members of the noumenal world. Following Kant's official response that appearances cannot define the concept of freedom (MM 6:226), I have likewise argued in the language of worlds that what we choose in the phenomenal world cannot legitimately explain how the will operates in the noumenal world. With this idea, I have shown that there is no contradiction

between the fact that we sometimes make wrong choices and the idea that the human will is always good. In other terms, for Kant, some unobservable, but thinkable part of us never ceases to be good no matter how terrifyingly we act in a moral sense.

The other intended contribution of the thesis to the literature has been offering an interpretation of the idea that evil is unintelligible. This idea has been mostly defended by Korsgaard (1996). Korsgaard is right that “how we go wrong” can only be explained by our phenomenal lives. My Kant reading has differed from Korsgaard in one significant aspect. I have argued that in Kant’s view, the human will is always good and that it should be treated as part of the noumenal world. While Korsgaard’s reading seems to preserve evil volition in humans, I have argued that an evil will is a diabolically evil will for Kant, which he rejects humans possess.

Korsgaard does not usually discuss the term *Willkür*, the power of choice, in the issue of evil. I have maintained in this thesis that the *Wille-Willkür* distinction is crucial to understanding one sense in which evil is unintelligible for Kant. From the intelligible substrate of free will, we are members of both the phenomenal world and the noumenal world. This means that insofar as the human will is intelligible as the source of actions in the world, not only the will is law-giving, but also *Willkür* is law-following. Why we would choose otherwise is unintelligible because to the extent it can be explained, we are members of the phenomenal world only. In simpler terms, for Kant, the question of why someone would choose to act in the wrong way finds an answer only when we turn to the world and its reasons as possible explanations.

Kant’s response to Reinhold is then that the human will never fails to give the moral law for our maxims. Yet, our choices may fail to align with the good human will. Why we fail cannot be explained on the basis of volition or freedom, which is a

fact about the noumenal world, of which our wills are a part. It can be explained only in relation to where we see evil, the phenomenal world.



REFERENCES

- Allison, H. E. (1990). *Kant's theory of freedom*. Cambridge University Press.
- Allison, H. E. (2020). *Kant's conception of freedom: A developmental and critical analysis*. Cambridge University Press.
- Beck, L. W. (1996). *A commentary on Kant's critique of practical reason*. University of Chicago Press.
- Kılınç, B. (2011). *Kant on chance and explanation*. In D. Dieks, W. González, S. Hartmann, T. Uebel, & M. Weber (Eds.), *Explanation, prediction, and confirmation (The philosophy of science in a European perspective, Vol. 2)*. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-1180-8_30
- Bojanowski, J. (2023). Reinhold on free will and moral obligation: A Kantian response. In J. Saunders (Ed.), *Freedom after Kant: From German Idealism to Ethics and the Self* (pp. 27–44). Bloomsbury Academic. <https://doi.org/10.5040/9781350187788.ch-2>
- Fantasia, F. (2020). *Zum Unterschied von Willen und Willkür*. In F. Fantasia & C. A. Meli (Eds.) *Dialectica Kantiana: Ragione ed effettività nella tarda filosofia di Kant*, pp. 167-216. Madrid: Alamanda.
- Fugate, Courtney D. (2012). On a supposed solution to the Reinhold/Sidgwick problem in Kant's metaphysics of morals. *European Journal of Philosophy*, 23(3), 349–373.
- Guyer, P. (2024). Freedom and immorality: Ulrich, Schmid, Reinhold, and Kant. In *Kant's impact on moral philosophy* (online ed.). Oxford University Press. <https://doi.org/10.1093/oso/9780199592456.003.0006>
- Kant, I. (2015). *Critique of practical reason* (M. Gregor, Trans.) (2nd ed.). Cambridge University Press.
- Kant, I. (1902–). *Gesammelte Schriften, Akademie-Ausgabe, Band 23, Vorerbeiten zur Metaphysik der Sitten*. Berlin: de Gruyter.
- Kant, I. (1997). *Lectures on Metaphysics*. (K. Ameriks & S. Naragon, Eds.). Cambridge: Cambridge University Press.
- Kant, I. (1996). Religion within the boundaries of mere reason (1793). In A. W. Wood & G. di Giovanni (Eds.), *Religion and rational theology* (pp. 39–216). Cambridge University Press.
- Kant, I. (1998). *Critique of pure reason* (P. Guyer & A. W. Wood, Eds.). Cambridge University Press.
- Kant, I. (2000). *Critique of the power of judgment* (E. Matthews, Trans., P. Guyer, Ed.). Cambridge University Press.

- Kant, I. (2004). *Immanuel Kant: Prolegomena to any future metaphysics: That will be able to come forward as science: With selections from the critique of pure reason* (G. Hatfield, Ed.) (2nd ed.). Cambridge University Press.
- Kant, I., & Wood, A. W. (1996). *The metaphysics of morals* (1797). In M. J. Gregor (Ed.), *Practical philosophy* (pp. 353–604). Cambridge University Press.
- Kant, I., & Wood, A. W. (1996). *Groundwork of the metaphysics of morals* (1785). In M. J. Gregor (Ed.), *Practical philosophy* (pp. 37–108). Cambridge University Press.
- Korsgaard, C. M. (1992). Creating the kingdom of ends: Reciprocity and responsibility in personal relations. *Philosophical Perspectives*, 6, 305–332. <https://doi.org/10.2307/2214250>
- Korsgaard, C. M. (1996). Morality as freedom. In *Creating the kingdom of ends* (pp. 159–181). Cambridge University Press.
- Noller, J. (2024). *The fate of choice: Freedom and imputability in Kant and his early successors*. Leiden/Boston: Brill. (Forthcoming).
- Reinhold, K. L. (1792/2022). *Some remarks on the concept of the freedom of the will, posed by I. Kant in the introduction to the Metaphysical Foundations of the Doctrine of Right*. In J. Noller & J. Walsh (Eds.), *Kant's early critics on freedom of the will* (pp. 238–249). Cambridge: Cambridge University Press.
- Schmid, C. C. E. (1790/2022). *Attempt at a moral philosophy*. In J. Noller & J. Walsh (Eds.), *Kant's early critics on freedom of the will* (pp. 72–82). Cambridge: Cambridge University Press.
- Ulrich, J. A. H. (1788/2022). *Eleutheriology or on freedom and necessity*. In J. Noller & J. Walsh (Eds.), *Kant's early critics on freedom of the will* (pp. 9–23). Cambridge: Cambridge University Press.
- Paton, H. J. (1946). *The categorical imperative: A study in Kant's moral philosophy*. Hutchinson's University Library.
- Potter, J. (1974). Does Kant have two concepts of freedom?. In *International Kant Congress (4: 1974: Mainz)* & G. Funke (Ed.), *Akten des 4. Internationalen Kant-Kongresses: Mainz, 6.–10. April 1974, Teil 2: Sektionen 1,2* (pp. 590–596). De Gruyter. <https://doi.org/10.1515/9783110878707-069>
- Reinhold, K. L. (2008). *Briefe über die Kantische Philosophie Zweyter Band [Letters on the Kantian Philosophy Volume II]* (M. Bondeli, Ed.). Schwabe. (Original work published 1792).
- Sidgwick, H. (1888). The Kantian conception of free will. *Mind*, 13(51), 405–412. <http://www.jstor.org/stable/2247440>
- Silber, J. (1960). The ethical significance of Kant's religion. In *Religion within the limits of reason alone*. Harper Torchbooks.