KANT'S CONCEPTION OF CONSCIENCE AND REFLEXIVE JUDGMENT

UMUT ELDEM

BOĞAZİÇI UNİVERSITY

2018

KANT'S CONCEPTION OF CONSCIENCE AND REFLEXIVE JUDGMENT

Thesis submitted to the

Institute of Graduate Studies in Social Sciences

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Philosophy

by

Umut Eldem

Boğaziçi University

2018

DECLARATION OF ORIGINALITY

I, Umut Eldem, 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 08.01.2018

ABSTRACT

Kant's Conception of Conscience and Reflexive Judgment

This dissertation consists of an explanation and interpretation of Kant's concept of conscience. By a detailed examination of Kant's Critical writings, I aim to show what roles the concept of conscience plays in his moral theory. This requires analyzing certain moral and cognitive concepts and principles as they relate to the phenomenon of conscience. Kant's account of rational judgment, especially in its reflexive aspect, is of vital importance for this project, as is his conception of the moral law, freedom of the will and our form of sensibly affected rational agency. After a general overview of Kant's philosophy which explains these concepts, I look at two important sources in order to reconstruct Kant's conception of conscience: the *Lecture Notes on Ethics* and the *Metaphysics of Morals*. Afterwards, I focus on issues that are related to our everyday understanding of the conscience, such as religious convictions, personal integrity, moral imputation and sincerity. The culmination of the dissertation consists in a discussion of Kant's ideas on moral education and how conscience could be cultivated.

The literature on Kant's conception of conscience has not explored its intricate relations with sensibility, reflexive judgment and religion. My thesis is an attempt to fill this gap. The elaboration of the concept of conscience, along with the explication of the concept of moral character, can bring together Kant's various text on moral philosophy into a systematic unity, which would be attuned to a comprehensive moral education.

ÖZET

Kant'ın Vicdan Anlayışı ve Refleksif Yargı

Bu doktora tezi Kant'ın vicdan anlayışının açıklanması ve yorumlanmasından müteşekkildir. Kant'ın eleştirel döneminde yazdığı yapıtların detaylıca incelenmesi yoluyla, vicdan kavramının Kant'ın ahlak kuramında hangi işlevleri yerine getirdiğini göstermeyi hedefliyorum. Bu inceleme, vicdanla ilgili bazı ahlaki ve bilişsel kavram ve prensiplerin de analiz edilmesini gerektirmektedir. Kant'ın rasyonel yargı kavramı, özellikle de refleksif (kendisine-dönük) yanıyla, bu çalışma için hayati önem taşımaktadır. Ayriyeten Kant'ın ahlak yasası, özgür irade ve rasyonel özne kavramları da bu tartışmaya katkıda bulunacaktır. Kant felsefesine dair genel bir özet sunduktan sonra, vicdan anlayışını ele almak için özellikle iki önemli metne odaklanmaktayım: Ahlak Üzerine Ders Notları ve Ahlak'ın Metafiziği. Daha sonra, günümüzdeki vicdan anlayışıyla alakalı olarak Kant'ın vicdanla dinî inanç, kişisel bütünlük, ahlakî sorumluluk ve dürüstlük arasında kurduğu bağlantıları inceliyorum. Son olarak da bu bağlantılarla beraber Kant'ın ahlakî eğitime dair fikirlerini tartışmaya açıyorum.

Kant literatüründeki tartışmalarda, Kant'ın vicdan kavramının hissiyatla, refleksif yargıyla ve dinle olan ilişkisi henüz detaylıca incelenmemiştir. Tezimin amaçlarından birisi bu boşluğu doldurmaktır. Kant'ın vicdan kavramı, ahlakî karakter anlayışıyla beraber ele alındığı takdirde, Kant'ın ahlak kuramını tutarlı bir biçimde bir araya getirilebilecek niteliktedir. Bu bütünlük ise ahlakî eğitim için sağlam bir zemin oluşturacaktır.

CURRICULUM VITAE

NAME: Umut Eldem

DEGREES AWARDED

PhD in Philosophy, 2018, Boğaziçi University

MA in Philosophy, 2013, Boğaziçi University

BS in Physics, 2009, Gebze Technical University

AREAS OF SPECIAL INTEREST

Kant's moral philosophy, history of philosophy, ancient philosophy, social and political philosophy, Hegel's political philosophy, Heidegger

PROFESSIONAL EXPERIENCE

Project Assistant, Boğaziçi BAP, "Transcendental Critique and Pragmatic Realism: Hegel's Reconstruction of Kant's System of Cognitive Principles" with Prof. Kenneth Westphal, (project number: 10081), 2015-2017

Instructor for Social Sciences, Piri Reis University, 2014-2015

English Teacher, B&K Foreign Language Education Center, 2009-2010

PUBLICATIONS

Journal Articles

Eldem, U. (2017), "The Concept of Conscience in Kant and Heidegger", *Felsefi Düşün*, 8, 178-201

Book Reviews

Eldem, U., (2018), "Kant on Conscience", forthcoming, *Studi Kantiani*

Book Chapters

Eldem, U., (2017a) “Heidegger on Language, Thought and the Human Being”, Gerhard Thonhauser (ed.), *Perspectiven mit Heidegger*, Alber Philosophie, 252-263

Theses

Eldem, U. (2013), *Descartes’ Understanding of Free Will*, Boğaziçi University

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor, Kenneth Westphal, for his excellent guidance and support. This dissertation would not have been written without his valuable comments, criticism and suggestions. It has certainly been a privilege to work with him. I would also like to thank the other members of the jury, Chryssi Sidiroupoulou, Yıldız Silier, Türker Armaner and Kaan Ökten, both for taking the time and effort to evaluate my dissertation and for instructing and teaching me on various occasions. In this regard, I would also like to thank other members of the department of philosophy in Boğaziçi for their guidance and instruction throughout the years: Murat Baç, Johannes Fritsche, Pınar Canevi, Stephen Voss, İlhan İnan and Lucas Thorpe. A special thanks is also owed to the editor at SBE, Peggy Alptekin.

I have also received invaluable insights, comments and, most importantly, encouragement and inspiration through hours of discussion and conversation from my dear friends: Gözde Yıldırım, Seniye Tilev, Çağlar Çömez, Hüseyin Kuyumcuoğlu, Elif Karakaya, Cansu Akarsu, Gizem Atalay, Fatma Kaya, Merve Tapınç, Nazım Gökel; Can Özbey and Melih Barsbey. You have all truly been the best part of the last seven years I have spent at Boğaziçi.

This dissertation was supported by the BAP Commission of Boğaziçi University for a duration of two years (project number: 10081).

Last but not least, I would like to thank my parents, Nesrin and Vasfi Eldem for their endless support. Ultimately, I owe them everything. I would also like to express my gratitude to my family (both current and prospective!) for their hospitality. A special thanks is also owed to Mümtaz Akşura for being an extremely understanding and supportive roommate for three years.

For Sena Pasin,

my soul, my inspiration, my heart

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION 1

CHAPTER 2: GENERAL OVERVIEW OF KANT'S MORAL PHILOSOPHY 12

2.1 Reason and judgment 15

2.2 Moral law, maxims and freedom 21

2.3 Moral anthropology 38

2.4 Reflexive judgment and moral feelings 40

CHAPTER 3: KANT'S CONCEPTION OF CONSCIENCE 51

3.1 Conscience in the Vigilantius Lecture Notes (1793-4) 52

3.2 Conscience in the Metaphysics of Morals (1797) 60

3.3 The infallibility of conscience 71

3.4 Secondary literature on the infallibility of conscience 80

CHAPTER 4: CONSCIENCE AND RELIGION 87

4.1 Conscience in the Theodicy essay (1791) 87

4.2 Divine commands and the highest good 94

4.3 Conscience in the Religion (1793) 105

CHAPTER 5: CONSCIENCE AND MORAL EDUCATION 122

5.1 Conscience and critique 127

5.2 Methods in moral education 132

CHAPTER 6: CONCLUSION 144

LIST OF ABBREVIATIONS

All references to Kant's works are from The Cambridge Edition of the Works of Immanuel Kant, unless stated otherwise. All details of publication are listed in the section entitled References. The year indicates the compilation the work belongs.

A/B	The Critique of Pure Reason (1998)
CL	Collins Lecture Notes on Ethics (1997)
G	Groundwork for the Metaphysics of Morals (1996)
HL	Herder Lecture Notes on Ethics (1997)
KpV	The Critique of Practical Reason (1996)
KU	The Critique of the Power of Judgment (1998)
ML	Mrongovius Lecture Notes on Ethics (1997)
LA	Lectures on anthropology (2012)
MS	The Metaphysics of Morals (1996)
NF	Notes and fragments (2005)
O.	What does it mean to orient oneself in thinking? (1996)b
Rel.	Religion within the Boundaries of Reason Alone (1996)b
T	On The miscarriage of all philosophical trials in theodicy (1996)b
VL	Vigilantius Lecture Notes on Ethics (1997)

CHAPTER 1

INTRODUCTION

This dissertation explicates and reinterprets Kant's concept of conscience, and its place within his moral philosophy. I aim to show what roles conscience plays in our moral lives, by examining its relations with different capacities and activities of the mind as Kant conceived them. So doing requires investigating the details of Kant's account of rational judgment, especially its reflexive aspect, together with his theory of moral agency, aesthetic predispositions, moral religion and anthropology. I shall also address issues that are related to our everyday understanding of the concept, such as religious convictions, personal integrity, imputation, sincerity and self-reflection. The culmination of the dissertation will consist in a discussion of Kant's ideas on moral education

The literature on Kant's conception of conscience has not explored its intricate relations with sensibility, reflexive judgment and religion. My thesis is in part an attempt to fill this gap. I claim that the elaboration of the concept of conscience, along with the explication of the concept of moral character and its development can bring together Kant's various text on moral philosophy into a systematic unity, which would be attuned to a comprehensive moral education.

It will be helpful for our purposes to put forward some preliminary remarks about what is required of any conception of conscience, which can adequately reflect our common sense understanding of the concept. No matter what our moral theories or preferences are, it is clear that our concept of conscience should contain a sense of moral awareness, especially in difficult moral dilemmas or clashes between different moral

principles. In a given moral conflict, the call of conscience is a call that brings our moral constitution to bear on the situation at hand. Should we fail to live up to our own moral commitments in a particular situation, where we could have done better, we are bound to feel the pangs of conscience. This is part of what is understood by our everyday experience of conscience.

At the same time, there is a sense in which conscience relates to our moral character as a whole. In the *Universal Declaration of Human Rights* it is stated that all human beings are endowed with reason and conscience, and that everyone has the right to a free conscience. Having a free conscience is usually associated with having the ability to choose or to change one's religious beliefs freely, and not being subject to coercion in these matters, but perhaps this freedom could be extended to include other kinds of moral commitments, such as vegetarianism, or ethical stances about biomedical issues such as genetic research, euthanasia, abortion and so on. No matter what we may think of the intricate and complex structure of our moral constitution, we can not deny that conscience is a part of our self-conception, and due to this fact, we have formulated a right (or a set of rights) which has universal scope. In this respect, then, it is safe to say that there is both a general and a particular aspect of conscience.

Since we associate conscience with a specific kind of right, the concept of conscience espoused here must also be politically significant, if it has any chance of being used in our day and age. The obvious example of the political significance of conscience is reflected in the phenomenon of conscientious objection to mandatory military service. This aspect of conscience is directly tied to personal integrity. In principle, conscientious objectors refuse any demand from the state which would

disintegrate their moral constitution by ordering them to violate one of their deepest moral commitments. From this perspective, conscientious objectors argue that states should have no right to force someone to act in a way which the person regards as contrary to his/her moral beliefs.

The political import of conscience can also be grasped via certain exemplars from recent (as well as ancient) history. These are usually people who have stood against what they took to be the unjust status quo, relying not upon the generally accepted political or moral customs, but rather acting in defiance of them for a better and more just society. We remember these figures in a positive light and we accept them as having been vindicated in history. We laud Socrates for his unwavering commitment to his own principles in the face of prosecution. Rosa Parks and Martin Luther King are regarded as champions of conscience, rising up to the challenge of disobeying segregationist laws in the United States. Mahatma Gandhi and Nelson Mandela could also be counted among such figures for the political struggles they lead in their countries.

That being said, defying the status quo by itself hardly calls for moral merit, one also has to act according to the correct moral principles and actually show how and why the status quo ought to be challenged and changed. We do not commemorate anyone who stood up against the established regime, we tend to respect those figures who we take to have acted from the correct moral principles. In fact, conscience could lead someone to commit actions that are regarded by many repulsive. As Paul Strohm (2011) puts it:

> [Conscience] can justify generous self-sacrifice, but selfish individualism as well. It can motivate an act of charity or an act of terror. The dictates

> of conscience can be Christian or pagan, divinely based or resolutely secular, selfishly nationalistic or generously international. (p. 2)

In a certain sense, then, conscience seems to be neutral with regards to the 'objective aspect' of moral situations, but is rather directed inward, toward our commitments and principles that we hold dear. One aspect of Kant's account of conscience will also reflect this quality, as it will turn out that conscience is intimately connected with our sense of personal integrity.

Now, what is characteristic of the "champions of conscience" is that they gave rise to a certain tension in their societies by drawing attention to the unjust laws, institutions or customs that govern them, whether implicitly or explicitly. They chose not to go along with the general opinion and instead pointed out how and why these customs and laws needed to change. In his Letter from Birmingham Jail, Martin Luther King (1963) refers to Socrates in this connection:

> Just as Socrates felt that it was necessary to create a tension in the mind so that individuals could rise from the bondage of myths and half truths to the unfettered realm of creative analysis and objective appraisal, so must we see the need for nonviolent gadflies to create the kind of tension in society that will help men rise from the dark depths of prejudice and racism to the majestic heights of understanding and brotherhood.

It seems plausible that a certain kind of tension in the mind (and in the community) is needed to conduct the individual (s) into a better moral order and understanding. This tension can be brought into the open by putting forward an ideal and showing how we fail to live up to it. This ideal need not be something entirely novel, it will typically be an ideal already embedded in our everyday understanding of morality or justice. The key

point is to ensure that people start thinking for themselves, rather than relying on habitual and general opinion, so that some sort of change may begin.

In the same text, we come across two important features for procuring this tension. The first is to disobey the law that we take to be unjust, but to do so "openly, lovingly and with a willingness to accept the penalty", just as Socrates did. This act of disobedience can lead to the questioning of the law itself, since it might urge other people to disobey too, on the grounds that it is unjust. What was once a general rule or custom can, in this way, be called into question. The second is to expose injustice "with all the tension that its exposure creates" to the "light of human conscience".

Conscience, then, is related to a certain kind of tension and a willingness to expose injustice. However, the tension need not arise between society and the individual, but could also arise within the person. Conscience can be directed to oneself as well as to the society at large. In both instances what is required is being sincere and truthful, both with oneself and with others. This attitude of sincerity, I shall demonstrate, is one of the basic aspects of Kant's account of conscience.

We incorporate a certain moral ideal into our conscientious thought to compare and contrast our own conduct with, both individually and collectively. I shall try to show that Kant's moral philosophy is well-suited for such an understanding of conscience, and can provide a sound interpretation of why and how we tend to regard the aforementioned historical figures as exemplars of conscience. What is more, Kant's theory of education may provide a method for raising conscientious people in general.

How, then, can we create this tension in our minds which can lead to our moral improvement? Answering this question requires investigating what our moral character is like. Throughout my discussion, I shall keep the concept of moral character as wide as possible, to include such features of our personality as: temperament, religious beliefs, aesthetic preferences, moral and political commitments as well as ideals, rules of conduct, behavior and so on.

Also included in our moral character must be something like the way in which we make sense of ourselves, our self-conception. For instance, religions in general could be construed as ways of making sense of our existence as a whole–however correctly or incorrectly. Furthermore, it should include at least a general understanding and recollection of our past (morally or practically significant) choices–as we shall see, this is where reflexive judgment would be useful in constructing a rough systematization of our past actions and prospects for future conduct. My general idea here is that the nagging of conscience is a direct and genuine response of our moral character, applied to various domains like politics, religion, morality and even the arts.

How do we come to have a moral character? According to Kant, human beings are the only kind of beings (that we know of) which require education (P 9:441) and they are nothing except what education makes of them (P 9:442). I shall explore Kant's ideas of moral education and their relation with his conception of conscience in the fifth chapter. For now, it will suffice to say that the formation of character (*Gesinnung*) is what Kant understands as the aim of moral education. Our character is the result of our moral education, which we typically receive from our family, our schools and our society.

In his book entitled *Voices of Conscience*, Thomas Green (1999) defines moral education as the process by which we "cultivate reflexive judgment on our conduct and character" which results in the "formation of conscience" (p. 61). Unpacking these ideas about character, conscience, education and reflexive judgment is the general aim of this dissertation.

Before going further with Kant's conception of conscience, it will be helpful to look at some prevalent notions of conscience that have been discussed in the history of thought. In this brief discussion I shall rely mostly on Richard Sorabji's comprehensive book on the history of the concept of conscience, entitled *Moral Conscience Through the Ages*. According to Sorabji (2014), the first instances of the notion of conscience that appear in Ancient Greek texts can be translated literally as: "sharing knowledge with oneself of a defect" (*suneidenai*), the defect usually being one of moral misconduct (p. 12). What is essential in this notion is that it consists of a reflexive act: that is, the agent turns toward herself and assesses her own conduct.

This mental act presupposes that the agent has a general idea of what a moral wrong consists in. It further requires the ability to examine oneself from the perspective of another person. As we shall see, these features of conscience have remained with considerable stability throughout the ages. Kant's account of conscience also incorporates this basic feature of *suneidenai* in that it is primarily an affair that an agent experiences within herself.

The experience of conscience entails a certain kind of self-scrutiny or self-interrogation. Sorabji (2014) notes that there were some mental exercises of self-scrutiny

commonly practiced among the Stoics, usually in the evenings with a calm and clear head. These practices involved giving an account of the day that has passed by asking a series of questions like: "what defect of yours have you cured today? What vice have you resisted? In what respect are you better?" (p. 28). These questions seem reminiscent of the way in which we experience conscience today.

As we shall see, Kant characterizes conscience as the "consciousness of an internal court" where "our thoughts accuse or excuse one another" (MS 6:438). This very phrase is present in Paul's letter to Romans (2:15). It seems the self-interrogation involved in the mental practices of the Stoics have survived for hundreds of years, as Kant has also picked up on them. It could plausibly be asserted that this aspect of character is central to our understanding of conscience.

In order to get a better sense of Kant's conception of conscience, we also need to look at its Christian interpretation. Sorabji (2014) explains a crucial tenet of the Christian understanding of conscience in this way:

> As with the original Greek concept, conscience in Paul is not the ultimate source of our knowledge of right and wrong. For him, the law in our hearts is the source. But conscience reveals our possession of that law. This makes the relationship between conscience and the inner law close, but not identity. (p. 31)

Furthermore, conscience for Paul is not "the voice of God", but rather "the agency that will give witness to God, as to whether one has lived by the inner law provided by God" (p. 34).

While Kant does not indicate that he adopted this view from the New Testament, this point is crucial for understanding Kant's account of conscience. As we shall see, Kant's understanding of conscience also relegates its role to a kind of consciousness of the law, as a "witness for the prosecution" so to speak, rather than characterizing it as the source of our knowledge of right and wrong. This distinction will be the basis for Kant's notorious claim that an "erring conscience is an absurdity" (MS 6:401).

Conscience can be said to be prevalent in many cultures. Strohm (2011) cites how something analogous to conscience can be found in many different languages and cultures, from Arabic to Hebrew, from Chinese to Latin, from Russian to Hindu (p. 3). Turkish culture is also imbued with the notion of conscience, it is utilized both in moral and in political contexts.

This is perhaps the most significant aspect of conscience: it is a universal phenomenon. This is the reason why we have attributed certain rights in order to protect it. Conscience, then, is one of the common traits of humanity which has a direct connection with morality; even though the specific moral customs vary widely from across different cultures. I shall argue that Kant conceptualizes conscience in a cosmopolitan way, which is a view that leads to significant political implications.

Sorabji (2014) draws a parallel between Adam Smith's conception of conscience and the Ancient Greek notion of *suneidenai* (p. 174). In Smith's account we come across a division of the self into two persons and conceives it as an "ideal spectator". This reflexive feature resurfaces in Kant's understanding of conscience as well, since he characterizes conscience as a "scrutinizer of hearts" (MS 6:438).

After this brief excursion into the history of the concept of conscience, let me present an outline of my dissertation. In the second chapter entitled "General Overview of Kant's Moral Philosophy", I present some of the key concepts and principles that will be relevant for my discussion and interpretation of Kant's concept of conscience. It is impossible to discuss any aspect of Kant's moral philosophy without having a clear idea of the concepts and principles that he employs throughout his critical writings. Especially salient in this regard are the concepts of rational judgment (and its reflexive use), moral law, freedom and moral anthropology

In the third chapter I look at two texts in some detail where Kant espouses his understanding of conscience. The first text is from the *Lectures on Ethics,* especially the notes taken by Johann Friedrich Vigilantius during 1793-4.The second text is from the second division of the *Metaphysics of Morals*, which is called the *Doctrine of Virtue*. This is the central text in any discussion of Kant's conception of conscience. It is in this text that the 'internal court' metaphor is to be found. This chapter will present some of the basic features of Kant's conception of conscience, such as its relation to imputation, subjective certainty, the feeling of awe and sincerity. I also focus on Kant's peculiar claim that "an erring conscience is an absurdity" (MS 6:401). This claim has recently given rise to significant discussions in Kant literature.

I present the relationship between conscience and religion in the fourth chapter. There are two main texts that I use for this discussion, *Religion Within the Boundaries of Reason Alone* and *On the Miscarriage of All Philosophical Trials in Theodicy*. These are texts primarily concerned with principles of biblical exegesis. Kant's definition of religion as "the recognition of all our duties as divine commands" plays a prominent role

and it is intimately connected with the concept of conscience (Rel. 6:84). I also discuss Kant's interpretation of the *Book of Job* as well his notion of the highest good.

The last topic I take up is Kant's pedagogy which will be found in the fifth chapter. Related to this discussion is also Kant's writings on anthropology, as well as the sections entitled "Doctrine of Method" in his critical writings (especially the first two Critiques). I examine how the method of the critique is related to conscience and also some of Kant's valuable insights regarding moral education. In my conclusion I put forward why Kant's moral philosophy in general, and his conception of conscience in particular, is still viable and valuable today.

CHAPTER 2

GENERAL OVERVIEW OF KANT'S MORAL PHILOSOPHY

In order to get a better sense of Kant's moral philosophy, let us begin with how Kant sets out the different branches of philosophy. Even at this early stage it will be helpful to keep in mind some of the idiosyncratic ways in which Kant uses traditional terms. Towards the end of the *Critique of Pure Reason*, we come across Kant's definition of metaphysics; "the whole (true as well as apparent) philosophical cognition from pure reason in systematic interconnection", which is a name that can be given to "all of pure philosophy, including the critique" (A841/B869).[1]

Metaphysics is further divided according to the speculative and practical uses of pure reason. Here, Kant insists that morality is "the only lawfulness of actions which can be derived entirely a priori from principles" and hence "the metaphysics of morals is really the pure morality, which is not grounded on any anthropology (no empirical condition)" (A841/B870). It is important to keep this idiosyncratic way of using the term metaphysics in focus, in order to understand even the titles of two of Kant's major works on moral philosophy, namely, the *Groundwork to the Metaphysics of Morals* and the *Metaphysics of Morals*, as a priori investigations into the principles of morality. As we shall see, even though this metaphysics is not "grounded" on any anthropology, it certainly depends upon it for its application (G 4:388-9, G 4:412 and MS 6:217).

[1]This definition is markedly different than what we have come to understand by the concept of metaphysics. Nowadays it is more common to regard metaphysics as a branch of philosophy.

Throughout his critical writings, Kant is adamant to distinguish between the empirical aspects of cognition and the a priori aspects. The same distinction is also required for morality. The main reason for this strategy is to formulate and represent rational principles in their "purity". What does Kant mean by purity here? We can get a sense of its meaning in a passage found in his *Lectures on Ethics*[2] in which Kant distinguishes between "the principle of discrimination of our obligation and the principle its performance or execution" (ML 27:1422). Kant takes this distinction to be analogous to the distinction between a measuring-rod and mainspring. Via this terminology, we can separate questions of criteria and justification with regard to our rational principles from questions of application and use of those principles.

In this connection, what it means for the moral law to be pure is that it does not require any empirical fact for its normative status. Implicit in this distinction is also the presupposition that normativity can not be reduced to merely empirical facts. Throughout the thesis I shall present numerous ways in which this presupposition is at work in Kant's thought.

Any in-depth discussion of Kant's philosophy has to rely on a number of texts, since Kant himself always worked out his ideas in a "systematic interconnection". A lot of the concepts that Kant employs throughout his critical writings originate in the *Critique of Pure Reason*, which will be the first source to be discussed in this section. It is in this text that Kant first explicates his conception of judgment. Some of the central concepts and styles of argumentation employed in this work resonate throughout Kant's

[2] This quotation is taken from Louis Infield's translation of the Lectures on Ethics, dated 1980, since this section of the Mrongovius lecture notes were not included in the Cambridge edition All other quotations from these lectures are from the Cambridge edition, translated by Peter Heath.

critical philosophy and it is crucial to establish these concepts to initiate the discussion on conscience.

Broadly speaking, we can categorize Kant's writings on moral philosophy into two groups; the first group concentrates on the a priori principles of morality (the measuring-rods, so to speak) and they consist of two major texts written in the 1780s: *The Groundwork for the Metaphysics of Morals* and the *Critique of Practical Reason*. These works mainly concentrate on the "pure" aspects of morality, such as the concepts of freedom, will, moral law, autonomy, obligation and practical reason and how these concepts are interrelated.

The second group of writings tend to incorporate some empirical conditions through which these a priori principles and concepts are to be applied and enacted in our everyday lives (the mainsprings for moral action). Incidentally, these are major texts written in the 1790s, namely *Religion Within the Boundaries of Reason Alone*, *Anthropology from a Pragmatic Point of View* and the *Metaphysics of Morals*. It is interesting to note that, while Kant was always careful to distinguish what is a priori in morality from what is empirical, his last major work in moral philosophy eventually contains both elements. In other words, it contains certain aspects of an anthropology.[3]

My primary focus in this dissertation will be these latter writings, together with some of the notes taken during his lectures on ethics and pedagogy. It is generally agreed in the secondary literature that these notes are not as reliable as the material Kant published himself. However, some aspects of Kant's theory of conscience and moral

[3] The reason for this is that Kant contrasts a general metaphysics with a special metaphysics. See Westphal 2010, p110

education are only to be found in these notes, which is why they have to be taken into account.

My discussion will also draw upon the *Critique of the Power of Judgment* as I intend to incorporate Kant's conception of reflexive judgment into my discussion of conscience. Furthermore, some essays that Kant wrote during his critical period will also be relevant for my purposes. The most significant of these essays is *On The Miscarriage of All Philosophical Trials in Theodicy*, as it contains an elaborate depiction of conscience.

All of these texts merit extensive discussion and exegesis, as they have each given rise to an immense secondary literature in many languages. For this reason, my discussion will inevitably leave out many aspects of Kant's moral philosophy. Nevertheless, I will be working with a certain interpretation of Kant and I shall now give an overview of what I take to be the basic aspects of this interpretation.[4]

2.1 Reason and judgment

As is well known, Kant's first major work in critical philosophy, which mainly deals with the theoretical use of reason, is the *Critique of Pure Reason*. This complex work deals with our cognitive faculties in general; what functions they are able to perform, which elements of cognition they are responsible for, how they are constituted, together with how we are aware of their functions and how they ought and ought not to be used.

[4] The main proponents of this interpretation are Onora O'Neill, Barbara Herman and Kenneth Westphal. Typically they focus on the normative structure of rational judgment and present Kant's philosophy without relying on Transcendental Idealism.

At the beginning of this work Kant states that there are "two stems of human cognition" which are "sensibility and understanding". By the former, objects are given to us and by the latter they are thought (A15/B29, A50/B74). Sensibility is further defined as "the capacity (receptivity) to acquire representations through the way in which we are affected by objects" (A19/B33).

After his exposition of the elements belonging to sensibility (in the "Transcendental Aesthetic"), Kant delineates three "faculties of higher cognition", which are understanding, reason and the power of judgment, which are discussed in the "Transcendental Analytic" (A131/B170). The crucial difference between sensibility and understanding is that the former is not spontaneous while the latter is (B68). Spontaneity in this context can be taken to mean an activity which originates in the mind, in contrast to a receptivity that the mind becomes subjected to due to its sensible aspect (merely perceptively). Kant also puts this difference in terms of what is determining and what is determinable in cognition (B152). Generally speaking, the understanding provides the determining forms or rules that are then applied to a determinable manifold of intuitions in order to achieve cognition.

According to Kant, cognition is only possible through a synthesis of the receptivity (via sense/as determinable) and spontaneity (via understanding/as determining) (A97). At this point it is crucial not to confuse Kant's account of sensibility with the standard/empiricist "raw data" account of perception. Sensibility also involves a certain synthetic activity, through forms of intuition and the schemata, together with the power of imagination that provides the manifold of intuitions which is determinable by the rules and concepts of the understanding.

Via our understanding, we can cognize spatio-temporal objects in our surroundings. However, these specific cognitions have to be brought into a systematic unity, in order for us to have a unified experience of ourselves and our surroundings. This is the principle of reason (in its logical use): "to find the unconditioned for conditioned cognitions of the understanding, with which its unity will be completed" (A307/B364). Reason, then, is our cognitive capacity that seeks the unifying principles of our particular experiences and thoughts. However, there are certain limits to this project and one of the main concerns of the first *Critique* is to determine and explicate the nature and significance of these limitations.

Famously, Kant distinguishes between a theoretical and a practical use of reason. To these uses of reason correspond two different kinds of cognition: theoretical and practical. Theoretical cognition concerns what exists while practical cognition concerns what ought (or ought not) to exist through our actions (A633/B661). In the theoretical use of reason, seeking the unconditioned, which lies beyond the constraints of spatio-temporal experience, leads us to paralogisms (fallacies concerning the soul) and antinomies (conflicting theses concerning the cosmos, matter, forms of causality and God). Accordingly, the critique of reason in its theoretical use has only a negative task: to prevent us from venturing beyond the limits of experience in our cognitions (A795/ B823).

In its practical use, however, reason is warranted to postulate the existence of three ideas: freedom, immortality and God (A800/B828). We come across these three concepts in almost all of Kant's moral philosophy. As we shall see, the postulate of freedom is necessary in order for there to be any morality at all, whereas the postulates

of God and immortality are required only for the "necessary object of a will determined by the moral law" which is the "highest good" (KpV 5:4). There are significant discussions to be considered with respect to these three ideals. In my discussion I shall focus primarily on freedom and God.

The power of judgment plays an indispensable role in both theoretical and practical cognition. As Kant puts it, understanding is the faculty of rules and concepts and the power of judgment is "the faculty of subsuming under rules i.e., of determining whether something stands under a given rule or not" (A132/B171). What is striking about the power of judgment is that "it cannot be taught, but only practiced" (A133/B172). The reason for this is that, if there were further rules for subsuming something under rules, we would require another capacity for determining whether this is a correct instance of subsumption and we would end up with an infinite regress of rules.

Even though the power of judgment cannot be taught, we can still derive many principles for its application through the explication of a transcendental logic, which can provide precepts "to correct and secure the power of judgment in the use of the pure understanding through determinate rules" (A135/B174). Furthermore, the power of judgment can be cultivated, or sharpened, by the use of examples (A134/B173).[5] As we shall see, these core features of judgment are of utmost importance for Kant's philosophy as a whole.

[5] The use of examples in moral education will be discussed in the final chapter of the dissertation.

It is vital to note that the power of judgment is not only exercised for the cognition of objects, but it is also exercised in practical situations. In the theoretical case, the power of judgment subsumes an intuition (given to us by sensibility) under a concept (given by the understanding). In the practical case, a universal rule is applied to a possible action in order to determine whether it is permissible, obligatory or forbidden (KpV 5:68). What this universal rule is and how it relates to action will be investigated in the next section where I discuss the formation of maxims for actions.

In his paper entitled "Universal Moral Principles and Mother Wit", Kenneth Westphal puts forward three distinctive features of rational judgment that he derives from the first *Critique*: it is inherently "normative, self-critical and social/communicable" (Westphal 2017a, 7). In making judgments, we are not only concerned about the content of the judgment, but also in its veracity. To ensure the veracity of our judgments we need to engage in reflection, which Kant stresses is required by all judgments. Reflection is "the state of mind in which we first prepare ourselves to find out the subjective conditions under which we can arrive at concepts" (A260/B316). What this means is that, in making a proper judgment, I need to be diligent in examining how I am making that judgment. In other words, a concern for making the correct kind of judgment is integral to my use of the power of judgment. I shall demonstrate that this reflexive aspect of judgment is closely connected with Kant's account of conscience.

The use of the power of judgment has to be spontaneous, normative, self-critical and communicable. These requirements for rational judgment can also be discerned by what Kant puts forward as the three principles of rational thought, which conjointly provide feedback for rational judgment: a) thinking for oneself, b) thinking from the

standpoint of others and c) thinking consistently (AP 7:200, KU 5:294). What is important is that there are no priority relations between these principles.

The first principle is about originality: I cannot at all be an agent if all I do is repeat the judgments and actions of other people. I need to be using my own spontaneous cognitive capacities for arguing, assessing, judging, understanding, rejecting, agreeing and so on. This originality of thought and action (that I am the originator of my own thoughts and actions, instead of a conveyor of other people's thoughts and actions) is a necessary condition for being a mature agent.[6]

The third principle, to think consistently, is undoubtedly a condition for personal integrity and the intelligibility of thought. There needs to be some consistency among an agents' thoughts and actions across time in order for that agent (as well as others) to recognize him/her to be an agent at all. However, this consistency hardly suffices for accuracy or veracity–we could be consistently wrong.

This is why the second principle always needs to accompany our thinking. In the theoretical domain, my judgments must be open to the scrutiny of others, since our judgments are always fallible.[7] In the practical domain, it is only by thinking from the standpoint of others that I can avoid violating the rights of others and it is by this principle that I can conceive of myself as part of a community. Furthermore, consistency ensures that my judgments are intelligible to all others so that they too can follow, assess, criticize, adopt or refuse my reasoning. Unless and until my judgments are intelligible and can be rationally adopted by others, it is eligible for justification.

[6] See the beginning of Kant's essay entitled "An Answer to the Question: What Is Enlightenment?" for an emphasis on thinking for oneself.
[7] See Westphal 2011d, p.35

2.2 Moral law, maxims and freedom

As we have seen, Kant asserts that the metaphysics of morals is concerned with the lawfulness of actions which can be assessed via a priori principles. This is why, in the *Groundwork*, Kant states that "the metaphysics of morals has to examine the idea and the principles of a possible pure will and not the actions and conditions of human volition" (G 4:391).[8] To be sure, in all of Kant's writings on moral philosophy, we find several examples from our everyday lives that purport to demonstrate how these principles can and ought to guide our practice. However, the principles themselves do not require the empirical features of human beings for their justification.

Famously, the *Groundwork* opens with the assertion that the only thing "good without limitation" is a "good will" (G 4:393). What is interesting is that according to Kant, "a good will is not good because of what it effects or accomplishes" but rather it is "good in itself" (G 4:394). What this means is that the moral worth of actions does not lie in its consequences, but rather in the principle through which they are initiated.

From the point of view of the moral agent, the principle through which an action is initiated (or decided upon) is called a maxim. Kant defines maxim as the "subjective principle of volition", where the objective principle is the "practical law" (G 4:401fn). The content of a maxim incorporates an end to be attained, as well as the reasons for the adoption of that end, and the means by which the end is to be attained, whereas practical laws express the principles that ought to govern the formation of maxims.

[8] This is yet another instance where Kant distinguishes between the measuring-rod and the mainspring.

Now, just as there are principles that regulate our exercise of the power of judgment, there are also rules and norms that govern our faculty of volition. In willing an action, we strive to do all that is in our power in order to achieve an end that we have set for ourselves. In this sense, willing is different from mere wishing (G 4:394). According to Kant, the normative principles involved in volition arises in the same way it arose for the power of judgment and understanding. As he puts it:

> We can become aware of pure practical laws just as we are aware of pure theoretical principles, by attending to the necessity with which reason prescribes them to us and to the setting aside of all empirical conditions to which reason directs us. (KpV 5:30)

As we have seen, there are certain features of rational judgment which we ought to incorporate into our thinking. In a similar way, there are certain principles of volition (to be expressed through maxims and actions) that we ought to uphold if we are to act in accordance with the requirements of morality. In fact, as we shall see, these principles converge to a certain extent. Before getting into the discussion of the moral law, it will be helpful to get a clear sense of what maxims are and what role they play in morality.

According to Barbara Herman (2007), "The actual object of primary moral assessment is not an action, but an evaluative principle (a "maxim") that represents what the agent intends to do as she judges it to be in some sense good" (p. 323). Actions by themselves cannot show the evaluative principles behind them, since the same action can be done for various reasons or principles. Maxims, on the other hand, involve such principles and hence become the primary object for moral assessment. In this way, our practical judgments concerning the rightness or wrongness of actions are incorporated

and expressed through our maxims. This is not to say that actions are irrelevant to moral assessment. It is better to understand Kant as working with a more extended concept of an action, which includes not only the corporeal behavior, but the reasoning behind it; that is, the maxim which guides and gives rise to the action. Certain kinds of action (such as coercion, deceit, murder, theft) will be prohibited by the moral law, no matter what kind of reasoning went into it.

As with many of Kant's basic concepts, there is a long standing debate about how to understand the nature of maxims. Onora O'Neill (1985) claims that they are "underlying intentions by which we guide and control our more specific intentions" (p. 162). This view is challenged by both Henry Allison (1990) and Robert Louden (2011). They claim that Kant's examples for maxims in the *Groundwork* exhibit varying degrees of generality and therefore maxims cannot only be underlying intentions.

O'Neill's interpretation has merit on two counts, first, it refutes a superficial understanding of universalizability that arises because "relatively specific intentions often can be universalized without conceptual contradiction" (O'Neill 1985, 176). Second, it allows for diversity under generality, since "the ways in which maxims can be enacted or realized by means of acts done on specific intentions must vary with situation, tradition and culture" (O'Neill 1985, 183).

In any morally significant action, then, a maxim represents what the agent actually adopts as a principle of action, whereas the "moral law" represents what any agent "ought to adopt" as a principle.[9] Kant describes the will as consisting of a formal

[9] See Westphal, 2010 , p114:"Adopt" here means to be able to follow consistently the very same principle in thought or action on the same occasion as one proposes to act on that maxim."

and a material component; an a priori principle and a material incentive. An incentive is the subjective ground of desire (G 4:427). Adopting the a priori principle itself as an incentive amounts to an action "done from duty" (G 4:400). The good will is the will that incorporates the moral law "as an incentive" (KpV 5:79).

The moral law (formulated as the categorical imperative) is the supreme principle of volition and is perhaps the most central concept of Kant's moral philosophy. Let us now discuss this notion and the concepts related to it, such as duty, volition and autonomy. To begin with, let us take the first formulation of the categorical imperative: "act only in accordance with that maxim through which you can at the same time will that it become a universal law" (G 4:421).

Why is this formulation of the moral law called a categorical imperative? An imperative is a command of reason that is necessitating for volition (G 4:413). What this means is that, following certain rational principles is constitutive for volition. Kant thinks that, in rational volition, there can only be two kinds of imperative, namely hypothetical and categorical. As he puts it:

> Since every practical law represents a possible action as good and thus as necessary for a subject practically determinable by reason, all imperatives are formulae for the determination of action that is necessary in accordance with the principle of a will which is good in some way. Now, if the action would be good merely as a means to something else the imperative is hypothetical; if the action is represented as in itself good, hence as necessary in a will in itself conforming to reason, as its principle, then it is categorical. (G 4:414)

Any instance of volition requires hypothetical imperatives, as these are formulations that pertain to the means through which we aim to achieve the end that we have set for

ourselves. This imperative has to be present in our maxim in order for there to be an instance of rational volition. However, hypothetical imperatives are neutral with respect to the morality of our actions, they are simply rational recommendations for what we take to be the best course of action for reaching a specific end, and therefore they are contingent upon our actual setting of an end.

The categorical imperative, by contrast, is not contingent upon the ends we set for ourselves. It is binding on us, no matter which end we set for ourselves.[10] Maxims, being subjective, do not necessarily agree with the objective ground of reason by themselves, this is why the demands of reason "reach us only as an imperative" (MS 6:213). What reason demands is a certain constraint upon our free choice of maxims and ends, which can be either external or self-constraint. This concept of a constraint is called duty, as "the moral imperative makes this constraint known through the categorical nature of its pronouncement (the unconditional ought)" (MS 6:379). Here, unconditional means that the imperative does not depend upon any ends or choices that an agent happens to have.

It is significant that reason is the source of the principle of morality. Onora O'Neill (1989) argues that the categorical imperative is not only a principle of morality, but "of all activity that counts as reasoned". In this sense, the categorical imperative is not an algorithm, but a "fundamental strategy" which is the "principle of thinking and acting on principles that can (not 'do'!) hold for all" (p. 58). As we have seen before, rational judgments had to be "communicable" and hence, "social". The same requirement also holds for the maxims by which we act. They ought to be communicable

[10] For a contrast of maxims and laws, see also the *Metaphysics of Morals*, 6:389.

and adoptable by any and all agents who find themselves in similar circumstances. Crucially, the actual adoption of these principles of action is a different, more empirical matter of inquiry.

As is well known, the categorical imperative has a number of formulations. Without getting into the highly complicated issue of the relations between these formulations, I shall briefly point out what matters for my purposes. The first formula is the one previously discussed. The second formula refers to "humanity as an end in itself" and the third formula is concerned with the "realm of ends". While Kant asserts these are all formulations of the same law, he also claims that the formulations that follow the first one are different in a "subjectively practical" sense, in that they bring the formulations "closer to intuition" (G 4:436).

According to Paul Guyer (2000), this reference to intuition is significant, since there can be "no proof of the reality of any concept without an intuition". Accordingly, the different formulations have to work conjointly (p. 160). The first formulation gives us the form of the moral law, which suffices for a "strict method", but the second formulation provides the "matter" and the third formulation the "complete determination". This is reminiscent of the three principles of rational thought discussed in the previous section, the first formulation involves a first person perspective (thinking for oneself), the second formulation involves a third-person perspective (thinking from the standpoint of other), and the third formulation concerns the consistency and systematic integration of all maxims (thinking consistently).

With these principles in place, let us now investigate what role freedom plays in this picture of morality. Every action issues from a maxim and aims to accomplish a chosen end. The first important point in the discussion of freedom is that "no one can have an end without himself making the object of his choice into an end" and hence setting oneself an end is an "act of freedom" rather than an "effect of nature" (MS 6:385). Furthermore, other people can force me to perform actions, but they can never force me to set myself an end (MS 6:381).

One has to admit that, throughout his moral philosophy, Kant is deeply troubled by the problem of free will, since he argues, already in the first *Critique*, that there can not be any theoretical proof or empirical experience of free will (Bxxviii).[11] This is because he construes the entirety of nature as causally determined: the thesis that any and all actions and motions depend on a necessary antecedent cause (A228/B281). If universal causal determinism were true, we could not make sense of a free action. Kant wants to keep both universal causal determinism and freedom of action intact, and this leads him to posit Transcendental Idealism as the only solution to this problem (A491/B519). Freedom of action is relegated to the noumenal, understood as independent from conditions of space and time, whereas everything that remains within spatio-temporality is necessarily construed as part of a universal causal mechanism.

As Kenneth Westphal (2017) has shown, Kant's account of freedom need not be construed as belonging to a completely different realm of existence at all, since Kant never justifies the thesis of universal causal determinism, upon which the entire problem

[11] See the Third Antinomy for an extended discussion of the problem of freedom and universal determinism. (A445-451/B473- 480).

is predicated.[12] If universal causal determinism can not be asserted, we need not relegate freedom of action into a different realm.

That being said, there is a sense in which free action can be said to be intelligible. I use the term intelligible, in contrast to empirical, to point towards the normative aspects of our judgment and action. For instance, let us think of a game of chess. We could have all the empirical facts about a game of chess, the material makeup of the pieces, the dimensions of the chess board, the movements of the pieces and so on: but in order to recognize a game of chess as a game of chess, we need to be aware of the typical rules that pertain to chess, or at least be aware that there is such a game called chess which is played by two people opposing each other with the same pieces. Unless and until we know these normative features of a game of chess, we can not recognize that what we are perceiving is a game of chess at all.

This is the reason why Kant wants to keep the freedom of thought and action in place, even though he espouses universal causal determinism. If all we had were empirical facts, there would be no room for normativity. This is because empirical facts can only give us what is the case, but normativity concerns "what ought to be the case". We need to be able to think beyond what is the case, since what ought to be the case may not have pertained yet.[13]

[12] In the same paper, Westphal also shows that Kant did not need to justify universal causal determinism for the theoretical use of reason either. See Westphal 2017 fn3. For an extended discussion of these issues, see Westphal 2004, §61.

[13] This is yet another instance where the difference between the measuring-rod and mainspring becomes relevant. The measuring-rod is normative, the mainspring is generally speaking empirical. We need the measuring-rod to ascertain whether the mainspring is functioning appropriately.

Given this interpretation of the relation between empirical facts and normativity, I shall avoid getting into the discussion of free will as conceived by Kant through his transcendental idealism. Instead, what will be enough for my purposes is what Allison (1990) calls the Incorporation Thesis (p. 54, see also Rel. 6:36, 6:42, 6:44). The upshot is that conscience presupposes a kind of moral agency which is able to act for reasons.

This point is of vital importance for any Kant interpretation. Moral agency, according to Kant, cannot and should not be construed or explicated on causal terms. This is because normativity in both action and thought requires spontaneity and hence, freedom.[14] This is also why Kant insists on the purity of moral principles. As we have seen, these principles as presented by the different formulations of the moral law, only refer to the rational nature of agents. This is a quality which any and all moral agents enjoy (we typically do not hold irrational beings morally responsible), and hence the moral law binds universally.

Kant's formulations of the moral law, then, do not involve the more specific/empirical features of human agency, but rather only rational volition. This is why moral laws are to be derived from the concept of a rational being as such (G 4:412). That being said, the reason that the moral law is an "imperative" for us is due to our specific form of agency, which is sensibly affected but not determined:

> [. . .] it is clear that [the necessity of acting in accordance with the moral law] depends only on the subjective constitution of our practical faculty that the moral laws must be represented as commands (and the actions which are in accord with them as duties), and that reason expresses this necessity not through a be but through a should-be:[. . .] (KU 5:403)

[14] As we have seen, Kant construes the faculty of understanding, reason and judgment as functioning spontaneously.

Let us briefly return to the first *Critique* where Kant distinguishes between an animal will and a human will. The animal will (*arbitrium brutum*), according to Kant, is pathologically necessitated, whereas the human will (*arbitrium liberum*) is pathologically affected, but not necessitated (A534/B562). This means that we have a certain power over our sensible nature. This definition will turn out to be crucial for Kant's moral philosophy in general.

Now, what is important for Kant at this juncture is that "[t]he ground of obligation here must not be sought in the nature of the human being or in the circumstances of the world in which he is placed, but a priori simply in concepts of pure reason" (G 4:389). This is consistent with his definition of a metaphysics of morals. The criteria set out in the categorical imperative is abstract and universal on purpose. Kant's account strictly rules out any criteria that appeal to the contingent features of human beings such as desires, incentives, beliefs, commitments, natural endowments, temperament and so on. Nor does it regard race, religion, nationality, gender etc. as relevant features for moral justification. This is the sense in which the moral law is universal in its scope, and the reason why our maxims must also conform to this requirement of universality.

That being said, certain features of our agency do play a role in the kinds of duties that are commanded by the moral law. For instance, were we creatures who could only think aloud (were we not able to hide our thoughts) we would not have a duty not to lie, since that would be practically impossible.[15] It is only because we can hide our thoughts that we have a duty not to lie. Furthermore, were the planet which we inhabit

[15] For Kant's discussion of this kind of rational being, see (AP 7:332).

considerably larger, we could have the possibility of avoiding each other indefinitely, and there would not be a constraint to enter into a cosmopolitan unity with each other.

The justificatory ground of the moral law does not rely on contingent facts about the human being. It does rely, however, on a necessary fact, so to speak. That fact is that, even through our common/everyday understanding, we are aware of the requirements of morality and we have the ability to act according to those requirements (given the power we have over our sensible nature). This is the definition of a morally responsible human being.

Stating that the ground of obligation is not to be sought in the nature of human beings does not entail that this nature is insignificant for moral purposes. It is only through the empirical makeup of the human being and its environment that these moral criteria can be applied; that is, used to shape, assess and criticize our maxims and actions. This is why any metaphysics of morals needs a practical anthropology for its application (G 4:412).

One basic aspect of moral agency is that one is imputable; which means that one is the author of one's actions and is therefore responsible for the consequences, as well as the reasons for that action. It is of utmost importance that we are free in setting ourselves an end and formulating a maxim which reflects our principles of volition with regards to that end. As Kant puts it: "I can indeed be constrained by others to perform actions that are directed as means to an end, but I can never be constrained by others to have an end: only I myself can make something my end" (MS 6:381).

As we shall see, Kant states that imputability is closely connected with conscience. However, in this preliminary account, we have to locate the ground of imputability on the freedom of choice (*Willkür*). What is also involved in imputability is that one is "previously acquainted with the law by virtue of which an obligation rests on (one's actions)". Thus we get the definition of a deed: "an action is called a deed insofar as it comes under obligatory laws and hence insofar as the subject, in doing it, is considered in terms of the freedom of his choice" (MS 6:223).

Imputability stems from free choice, as this capacity is the locus of the formation of maxims. As Kant describes it, all lawgiving incorporates two elements, namely a law and an incentive "which connects a ground for determining choice to an action subjectively with the representation of the law" (MS 6:218). What it means for me to give myself the law is that I turn my duty itself into an incentive; that is, I perform my duty only because it is a duty. This requires recognizing the moral law as authoritative and that therefore I am obligated to act in accordance with it. Autonomy is the condition in which I act by incorporating the incentive of duty into my maxim. Through the same power of free choice, I can incorporate some other incentive, in which case my action would become heteronomous. The key point is that no incentive can lead me to any action unless I take it to be a reason for action: this is called the incorporation thesis (Allison, 1990, 39).

The recognition that we are bound by the law results in the recognition of duties– the law as it is applied to our circumstances. Kant argues that duties toward ourselves are necessary in order for there to be any duty towards others. This is because "the law by virtue of which I regard myself as being under obligation proceeds in every case from

my own practical reason; and in being constrained by my own reason, I am also the one constraining myself" (MS 6:418).

This is one of the most important passages in the entirety of the Doctrine of Virtue.[16] It can be made more accessible if we take into account Henry Allison's interpretation of the difference between *Wille* and *Willkür*. To put it briefly, Allison (1990) shows that the term *Wille* designates both the legislative aspect of the will, and the will as a totality; that is, a combination of the legislative aspect with the executive aspect (*Willkür*) (p. 129).

Unless this point is kept in mind, the distinction between *Wille* and *Willkür* can be confusing. In the *Metaphysics of Morals* (as well as in the *Groundwork*), the will is sometimes taken as "practical reason itself" and it can "determine choice" (MS 6:213). From the first person perspective, we are aware of our will as having both a legislative aspect and an executive aspect. The legislative aspect is designated as practical reason, but the executive aspect is the part of will which plays the central role in the formulation of maxims: it is the agent proper. This means that we could either choose to act under the legislation of reason, or we could choose to act from other considerations due to this executive branch. Autonomy, then, could be construed as the condition in which the legislative aspect itself also becomes the executive aspect of the will.

Were I not aware of these aspects of my will, I would not be able to comprehend my being under any obligation at all. The consciousness of my own will, as both free and obligated, provides a model from which I recognize that I am obligated toward other

[16] See also MS 6:410 "Viewed in terms of its formal principle, ethics is the science of how one is under obligation without regard for any possible external lawgiving".

people in general, as a human being living under laws. To put it differently; it is only on the basis of the recognition of our capacity for self-constraint (in our will) that we are able to make sense of being obligated at all. This is possible only by being able to regard myself as imputable for my actions and this imputability is central to Kant's conception of conscience.

As Kant puts it: "Only freedom in relation to the internal lawgiving of reason is really an ability; the possibility of deviating from it is an inability*"* (MS 6:227). This passage can be taken to mean that "only actions from duty are free actions". However, this is not the issue in this passage; Kant is here concerned with the "expository principle" of free action, rather than what defines free action. His worry is this: if we start out from *Willkür* as the ground for *Wille*, this would result in a hybrid definition that leads to a morality of heteronomy, whereas what is needed is a pure definition which preserves the possibility of autonomous action.

What is wrong with hybrid definitions and heteronomous theories? Onora O'Neill (1985) explains why criteria for moral acceptability must not rely on natural phenomena, such as desires and inclinations, since, if we do: "there will be no types of act which would not be rendered morally acceptable by some changes in human desires" (p. 159). In other words, our criteria must not be susceptible to change by temporary empirical features of agents such as ends, desires or choices. Accordingly, our conception of moral agency must be such that "it has to work with a conception of action which has the sort of formal structure which can meet (or fail to meet) standards of consistency*"* (*ibid*. 161).

This formal structure can be reflected in maxims, since maxims incorporate our reasons for action. What gives rise to these reasons could be varied. Most typically, we could have certain inclinations that direct us to pursue the ends which would satisfy them. Part of what morality requires is that the pursuit of our ends does not violate the ability of other agents to pursue their ends.[17] This may often require regulating or constraining our inclinations (and hence our maxims and actions) due to our regard for the moral law. This is the most general way in which the concept of duty is understood- the external or self-constraint of inclinations (and actions arising from them) by the moral law (MS 6:379).

The division of the *Metaphysics of Morals* into the *Doctrine of Right* and the *Doctrine of Virtue* reflects the division between different kinds of constraint. Generally speaking the constraint involved in duties of justice is external constraint. The main objective in this kind of constraint is to ensure that the pursuit of ends in a society do not conflict with each other. Actions that hinder or outright destroy another agent's capacity to act rationally (such as deception, coercion or physical harm) are strictly forbidden. This prohibition can be enforced through legal means. As Kant puts it: "Strict right rests instead on the principle of its being possible to use external constraint that can coexist with the freedom of everyone in accordance with universal laws" (MS 6:232).

The doctrine of virtue, by contrast, is concerned with issues regarding self-constraint, and it is within this aspect of morality that conscience plays its primary role. This part of the *Metaphysics of Morals* is concerned with the kind of dispositions that an

[17] The formula of a realm of ends is the formula in which this requirement is pronounced most clearly. What is of utmost importance is that the various ends of human beings can co-exist in a systematic unity under a rational principle (G 4:433).

agent ought to have as well as the kinds of ends that ought be set for all moral agents. According to Kant, we have a wide duty of "assessing the worth of one's actions not by their legality alone but also by their morality (one's disposition)" (MS 6:393). The point is that moral motivation is an integral part of moral action. In doing the right action, one must also have the right motivation that initiates that action. As Barbara Herman (1998) puts it "actions that are only accidentally right, even if reliably so, are not responsive to moral concerns" (p. 256). What is required for morality, then, is the right kind of concern with morality that is somehow reflected in the action.

This is precisely the heart of Kantian ethics: You ought to do the right thing, because it is the right thing to do. Acting from respect for duty (*aus Pflicht*) is not the same as acting in accordance with it, which will be some form of action from inclination (*aus Neigung*) (G 4:398). The difference is immense; it is the difference between autonomy and heteronomy. That being said, actions done from inclination could also be the right action, in the sense that they would be legally correct. The right action can be imposed through legal means, but having the right disposition is a more subjective matter and hence there can not be any external laws that command to have a certain disposition.

What Kant means by autonomy is that we hold ourselves to the standard of the moral law, and expect others to take it as a standard as well. This is what makes our actions imputable to us, so it is what makes us morally responsible; we have the capacity to act in the right way for the right reasons. Automata could be programmed to act according to rules and principles. It is only through critical rationality that we are able to use, adopt, criticize, assess or dismiss principles. Furthermore, what is important is that

our endeavors in morality belong to us; that our moral improvement (or lack thereof) is imputable to us (MS 6:228). Kant defines imputation (in the moral sense) as "the judgment by which someone is regarded as the author of an action" (MS 6:227). This action is then called a "deed" (*factum* as Kant calls it in Latin).

Thus, from a Kantian perspective, mere compliance with the moral law will not suffice for the moral worth of actions. We need to act from the recognition of the authority of the moral law. To recognize the authority of the moral law (that is, respecting the moral law-and acting from that respect) is different from learning the content or the formulation of the moral law: it is about having the right disposition toward the moral law. As we shall see, conscience is ultimately related with having the right disposition toward the moral law. This disposition, I claim, will not only be intellectual, but it will also involve what Kant names moral feelings.

This is also why the formulation of the moral law is not enough for a complete moral theory; an in-depth discussion of moral character is necessary. Furthermore, we also require a discussion of the kind of society in which we live in. Throughout his corpus, Kant insists that it is not sufficient, from the moral perspective, that human beings only fulfill the requirements for a just society; to put it differently, compliance with the moral law in an external way (acting according to, or at least not acting as opposed to the moral law) is not enough for human beings to become moral (G 4:390).

Kant usually puts this point in terms of the "legality" of action and the "morality" of it (KpV 5:71, 5:81, MS 6:219, 6:392). Following the letter of the law is certainly integral to correct action, but unless it is accompanied with the right disposition, we can

not claim that any right action has moral worth. We need a kind of moral education which aims not only to teach the correct moral principles, but also to cultivate the pupil in order that s/he acquires the correct disposition toward those principles. In the following chapters, I shall demonstrate that Kant's conception of conscience can play an indispensable role in this latter task.

2.3 Moral anthropology

A more comprehensive account of conscience is elemental in establishing an understanding of the moral constitution of human beings. Only through this constitution can we apply the criteria inherent in the moral law to our everyday lives. This is part and parcel of what Kant called a "moral anthropology". Granted, he deemed this anthropology as "appendix" to his metaphysics of morals, in order to distinguish issues of criteria and justification from issues of application (G 4:388). However, both issues ought to be settled in order to establish a more comprehensive moral theory. Anthropology, then, ought to give an account of how human beings acquire a moral constitution through education (both in the family and in the society at large), and furthermore, what this constitution ought to be like in order for us to be morally adept at dealing with particular situations.

As we have seen, Kant's moral philosophy does not require transcendental idealism for its justification. However, it does need an account of how we come to form maxims, how moral principles can play a role in the governance of conduct and the formation of agency, how these abstract and universal principles can be brought closer

to the human form of cognition and sensibility. In short, it needs a specifically moral anthropology, which Kant never explicitly wrote. The main reason we need this anthropology is to make the moral law efficacious. I will show that conscience plays a central role in a moral anthropology that is amenable to his metaphysics of morals.

Throughout Kant's critical writings, we come across various references to moral anthropology. Already in the *Groundwork,* Kant asserts the need for such anthropology for the "application" of moral laws (G 4:412). In the *Metaphysics of Morals*, anthropology is conceived as "dealing with the "development, spreading and strengthening of moral principles" (MS 6:217).

Even though Kant repeatedly stresses the importance of moral anthropology, he is also concerned about the danger of mixing this anthropology with a metaphysics of morals, which would result in weaker principles. The reason for this is that the demands of morality are categorical, and hence they have to be universal and necessary. Indeed, the required universality can only be achieved by leaving the contingent aspects of human nature out of the justification of moral principles. Moral responsibility is attributable to us due to our rational capacities.

That being said, the *Metaphysics of Morals* itself is full of duties which can only pertain to beings like us, with our abilities and weaknesses, living in this planet which exhibits these properties. Thus, the actual *Metaphysics of Morals* inevitably contains parts of moral anthropology. Without a reliable account of our circumstances and our characteristics, reason can not be practical and the purity of the law can scarcely help.

2.4 Reflexive judgment and moral feelings

One important aspect of rational judgment to Kant's conception of conscience is the reflexive aspect of judgment. *The Critique of the Power of Judgment* is the main text in which this aspect of judgment is discussed, and it is to this text that I now turn.

In the First Introduction to the *Critique of the Power of Judgment*, Kant puts forward three kinds of judgment, each with its own a priori principles. These are theoretical, aesthetic and practical judgments. The latter two kinds are marked by purposiveness, aesthetic judgments relate to the feelings of pleasure and displeasure, and practical judgments "stand under the idea of a form of purposiveness that is qualified for universal law, as a determining ground of reason with regard to the faculty of desire" (FI 20:246). With regard to theoretical judgments, purposiveness can only play a regulative role with the "aim of a thoroughly interconnected experience" (KU 5:184). In case we do come across systematic unity among empirical laws, we are relieved of a need, that is, the need to be reminded that our attempts at an interconnected experience is accommodated by the unity in nature.

The key point is that we can only make sense of purposiveness via the reflexive use of judgments. Our intentional actions, that is, setting ends and striving to achieve them through the appropriate means, can not be understood merely mechanistically (by appeal to natural laws), since these actions and the thought processes that lead up to them are ought to be regarded as spontaneous and hence, free. Neither can certain organic formations in nature be understood merely mechanistically. If they were, the determining aspect of judgment would suffice. This is why we are "forced to think of

another principle . . . as the ground of the possibility of certain forms in nature" (KU 5:388).

This use of reflexive judgment is so central that even to have an "experiential cognition of the internal constitution of organized things", we require "the thought of a generation with an intention" (KU 5:398). Reflexive judgments are primarily about our state of mind (KU 5:264). What is especially significant for moral purposes is that, through the use of reflexive judgments, we are able to recognize purposive behavior which could only have arisen from an intelligent beings. As Kenneth Westphal explains:

> The Critical philosophy justifies the general principles required for this abductive inference, so when we observe behavior which cannot explained by natural causality, but which can only be understood as resulting from intentional purposiveness, we are entitled to ascribe sensibility, understanding and reason to the agent. (Westphal 2017, 12)

Reflexive judgments, then, are required in order for us to identify beings (including ourselves) which have rational capacities and hence can act according to reasons (good or bad). This rationality is the mark of humanity: it is what makes us morally responsible, which in turns bestows upon us certain rights and duties. It is clear from these points that reflexive judgments play a central role in all our rational undertakings, whether practical, theoretical or aesthetic.

Rudolph Makkreel (2002) ties reflexive judgment to our finitude: "Reflective judgment is our way of compensating for our finitude, without relying on a dogmatic faith in religious doctrines" (p. 215). Reflexivity, then, is also the way in which reason becomes the sole authority for itself. Even though reaching complete knowledge is

impossible, reflexive judgments allow us to foresee a comprehensive plan for the systematic expansion of knowledge (KU 5:386) as well as, I shall argue, the formation of a consistent character. In these endeavors, the normative principles at work in cognition and volition can be used reflexively to reach judgments that can warrant universal assent; since these principles are shared (or ought to be shared) by everyone.

This is the point that ties reflexive judgment with systematicity. No agent can have complete cognition of nature or society, whether from a religious, philosophical, political or scientific perspective. The search for truth in any domain has to be a joint effort, since all agents have certain cognitive limitations. However, this search can not be conducted by arbitrarily going through various principles, judgments or explorations, it must be grounded in certain principles which can be adopted universally by those who are conducting the search in question. The success of this undertaking is only possible by the intelligibility/communicability of our assertions, arguments and ideas. In this sense, asserting what we hold to be true must be both our right and our duty. In the following chapters I shall demonstrate that this specific right and duty is central for Kant's conception of conscience.

As Paul Guyer (2003) notes, this attempt at systematic unity presents us with an ideal "toward which we must always strive, but also [one] that we can only approximate and never fully attain" (p. 7). In the moral domain, this ideal corresponds to the realm-of-ends formulation of the moral law. Combining this moral ideal with the pursuit of happiness, which Kant thinks is naturally found in all human beings, ultimately leads to his conception of the highest good. As I shall demonstrate, moral ideals are of central importance in all of Kant's writings on moral philosophy and conscience.

For the time being, let me continue with reflexive judgments. I have stated that any theoretical and practical endeavor ultimately requires a joint effort. What we need, then, is a plan that is structured in a way which can be followed by all parties involved and the most suitable plan would be the one that springs from reason itself, and hence is universal. This is one of the core insights of Kantian philosophy, summarized in the following passage:

> [R]eason has insight only into what it itself produces according to its own design; that it must take the lead with principles for its judgments according to constant laws and compel nature to answer its questions, rather than letting nature guide its movements by keeping reason, as it were, in leading-strings; for otherwise accidental observations, made according to no previously designed plan, can never connect up into a necessary law, which is yet what reason seeks and requires. (Bxiii)

Keeping reason in leading-strings is to keep reason in immaturity by forcing it to rely on external authorities. Our best judgments have to be our own, springing from the spontaneous use of our cognitive powers, while the regulative principles of reason expressed through reflexive judgments ought to guide our theoretical and practical undertakings. It will turn out that, according to Kant, conscience is the witness to our best judgments, especially in the practical domain. By reflecting on our past decisions, actions and maxims through the experience of conscience, we can assess both our particular decisions and our character as a whole and seek to act in ways that are consistent and continuous with our character.

This consistency, however, will not enough for moral accuracy, hence we also require the feedback of our peers. The main reason for making consistent and intelligible

judgments is to allow for other to follow, assess, criticize, correct them on rational grounds. The constraint on intelligibility and communicability is, then, a *sine qua non* of any joint effort, whether in theoretical or practical undertakings.

Another significant feature of conscience will turn out to be its relation to feelings. Ascertaining the relationship between the power of judgment and feelings will be crucial for understanding and assessing Kant's account of conscience. For this reason, we need to take a closer look at our cognitive and affective constitution as Kant understood it. In a particularly dense passage, Kant explains the relation between judgment and feeling in this way:

> Now between the faculty of cognition and that of desire there is the feeling of pleasure, just as the power of judgment is contained between the understanding and reason. It is therefore to be suspected at least provisionally that the power of judgment likewise contains an a priori principle for itself, and, since pleasure or displeasure is necessarily combined with the faculty of desire (whether, as in the case of the lower faculty of desire, it precedes the principle of that faculty or, as in the case of the upper, it follows only from the determination of that faculty through the moral law), it will likewise effect a transition from the pure faculty of cognition, i.e., from the domain of the concepts of nature, to the domain of the concept of freedom, just as in its logical use it makes possible the transition from understanding to reason. (KU 5:178)

This passage warrants extensive discussion, since it is here that we see the unifying function of judgment in Kant's overall system. What I take Kant to mean in the distinction between "upper" and "lower" faculties of desire are precisely what he termed *Wille* and *Willkür* (KpV 5:22. In *Wille*, which is the legislative aspect of the will, the determination of the executive faculty through the moral law raises the feeling of respect, whereas in the lower faculty of *Willkür*, pleasure or displeasure can "precede the

principle": in other words, we can act by prioritizing our inclinations (or our self-love in general) over the demands of the moral law (resulting in what Kant calls in the *Religion*, evil). What, then, is the relation between pleasure and judgment?

The faculty of pleasure brings together the faculty of cognition and the faculty of desire; that is understanding and volition. In a similar vein, judgment is located between understanding and reason. What emerges from this picture is that understanding is the faculty by which we cognize particulars, however, there has to be some organizing principle by which these diverse cognitions are to be united under a systematic whole and this is done by reason through the use of reflexive judgment.

Judgment, in its determining role, aids the understanding by subsuming particulars under concepts. In its reflexive role, it seeks to unify diverse particulars under a systematic whole; so we return from the particular to the general by directing our attention to the concepts and principles that we possess. In this way, the determinate use of judgment can be understood as dogmatic, whereas the reflexive use of judgment can be understood as critical (KU 5:395).

In the *Critique of Judgment* we come across two specific kinds of pleasure which relate not to external objects (that which is agreeable) nor to a determinate concept (satisfaction in the good), but to our cognitive faculties via aesthetic judgments (KU 5:244). In the judgment of the beautiful, pleasure arises from the harmony of imagination and understanding in free play. In the judgment of the sublime, awe arises from the conflict between reason and imagination (KU 5:256). In the next chapter I shall argue that this latter sensation is central to Kant account of conscience.

Kant defines feelings in general as the "capacity to be susceptible to pleasure and pain" (MS 6:211).[18] Inclinations are "habitual sensual desires" and they are typically incorporated into our maxims as reasons for action (A 7:251). There is, however, another type of feeling (or set of feelings), associated with our engagement with the moral law. I shall consider the details of moral feelings in the next chapter when I discuss the "aesthetic predispositions" in the *Metaphysics of Morals*. In the third *Critique*, Kant claims that "our predisposition to the feeling for practical ideas" is a universal quality of all people who have a "healthy understanding" (KU 5:265). Following from this point and expanding upon the predisposition, Kant distinguishes between two kinds of feeling with respect to their relation to the moral law in the *Metaphysics of Morals*. As he puts it:

> Every determination of choice proceeds from the representation of a possible action to the deed through the feeling of pleasure or displeasure, taking an interest in the action or its effect. The state of feeling here (the way in which inner sense is affected) is either pathological or moral. The former is that feeling which precedes the representation of the law; the latter, that which can only follow from it. (MS 6:399)

Pathological feelings are typically associated with actions done from inclination (*aus Neigung*). Moral feelings, on the other hand, follow from the representation of the law. It seems clear in this passage that our engagement with moral situations involve both affection and intellect. Kant clearly states in many texts that the affective aspect of morals (which can be construed as a moral anthropology, since it relates specifically to human agency) is not a part of the metaphysics of morals, and hence plays no role in the

[18] In her paper entitled "Kant's Taxonomy of Emotions", Kelly Sorensen refers to some passages in the *Critique of the Power of Judgment* where Kant presents the importance of emotions for our experience. My discussion of emotions here follow upon Sorensen's exegesis (See Sorensen 2002 for further details).

justification of moral principles (G 4:388, 4:411; MS 6:217). Nevertheless, this aspect is a "subjective presentation" of morals" through which the moral law "makes its efficacy felt" (MS 6:406).

In her book entitled *Kant's Theory of Virtue*, Anne Margaret Baxley (2010) notes that the moral feeling is "the way in which we register the normative force" of moral considerations (p. 149). Moral feelings, then, ought to be cultivated so that we get used to recognizing and registering moral considerations more easily. The distinctive qualities associated with moral feelings can help us be more alert and attentive to the morally relevant features in our everyday lives. Kant was well aware of this point, as he wrote about "the feelings that accompany the constraining power of the law" such as "horror, disgust" which make "moral aversion sensible" (MS 6:406). These feelings can act as indicators of morally salient features in our surroundings.

Nancy Sherman (2014) describes the affective aspect of morality as giving rise to a qualitatively distinct experience, in which we "notice [the morally relevant features of situations] with a certain intensity or impact that would be absent if emotions weren't engaged" (p. 12). Presumably, if we were purely rational beings, we would not need our emotions to play this kind of role in matters moral. Kant repeatedly makes this point when he discusses what a pure will would be like (G 4:390, KpV 5:30, MS 6:213). Alas, we are creatures who have both sensibility and rationality i.e. we are *arbitriae liberae*.

Our engagement with morality, then, has to accord with our natural constitution, and this is what Kant aims to explore especially in his texts written during the 1790s. Moral feelings, according to Kant, are "implanted in us by nature to do what the

representation of duty alone might not accomplish" (MS 6:457). Morally attuned emotions play a role not only in the recognition of morally relevant features or in registering their normative force, but also in moral motivation. Important here is that moral feelings are a natural part of human agency, and hence they can not play a role in the justification of moral principles, they can not serve as measuring-rods but rather belong to the mainspring of human action. Feelings in general are not under our control, we can not utilize them spontaneously, and hence they belong to the receptive aspect of our minds (KpV 5:22).

This point, however, needs to be qualified if we are to make sense of how we can cultivate these feelings so that they aid us in our moral purposes. As we shall see, moral feelings have an intellectual aspect to them, and we do have some control over them. Otherwise it would not make sense for Kant to argue that we have a duty with regard to their cultivation (MS 6:457).

Now, given this distinction between what belongs to the receptive aspect of the mind (inclinations and feelings) and what belongs to its spontaneous powers (judgment and volition) we need to mark another important insight. In the first *Critique*, Kant presents his view of philosophy as "the legislation of human reason" which has two main objects: nature and freedom (A838/B866). Accordingly, his critical philosophy can be regarded as an attempt to bring these two main objects into a single philosophical system. Moral feelings seem to be one instance where intellectual and spontaneous activity is closely linked with our sensible nature. This link would be incredibly difficult to establish from a transcendental idealist perspective, since we could not make sense of a unification of noumena (spontaneous and free action of the mind) with phenomena

(receptivity). I believe this is one of the advantages of the Kant interpretation that I am working with.

As we have seen, according to Kant, our natural constitution involves sensibility, inclination and feelings. This natural constitution can be cultivated for purposes of morality and indeed, I shall demonstrate that conscience plays a central role in this endeavor. This point is especially significant for Kant, with regards to his concept of the highest good. If our natural predispositions can be cultivated so as to serve our moral aims, we could also make the regulative claim (via the reflexive use of judgment) that nature in general is suited for the realization of morality. This point leads to the discussion of the highest good and Kant's conception of God: an issue I take up in the fourth chapter.

In this way we can propose a world in which the satisfaction of our inclinations can coexist with morally right conduct. Via the concept of a purposiveness in nature, provided by the reflexive use of judgment, the highest good (the final end of practical reason) is acknowledged to be possible (KU 5:196). I shall discuss the notion of the highest good in the fourth chapter, as it is a central idea in moral education.

In this chapter I have put forward the central features and concepts of Kant's critical philosophy. I have shown the importance of distinguishing questions of criteria and justification from questions of application and use. I have also presented the particular interpretative framework within which this dissertation operates. This entailed some significant points about the moral law, freedom, moral anthropology and rational judgment (especially in its reflexive use). What we have now are the necessary tools for

understanding not only Kant's conception of conscience, but his critical philosophy in general. With these points in mind, let us now turn to Kant's discussion of conscience.

CHAPTER 3

KANT'S CONCEPTION OF CONSCIENCE

In my introductory chapter I have stated that conscience is one of the central concepts of our everyday moral practice. Now, what role does the concept of conscience play in Kant's moral philosophy in general? Andrea Esser (2013) argues that, in Kant's moral philosophy, "conscience is assigned neither a causal role nor a leading role in terms of content, nor a generally or systematically important role, but only a marginal one" (p. 281).

Esser claims that Kant developed a more restricted account of conscience so that it does not compete with the dictates of practical reason. This is certainly correct; since, as I shall demonstrate, conscience belongs to moral anthropology. This empirical part is required for the application and use of moral principles and not their content or justification. To continue the analogy made earlier, conscience is related to the mainspring and not the measuring-rod with regard to moral actions.

Contrary to Esser, however, I shall argue that the concept of conscience does play a systematically important role in Kant's moral philosophy. Conscience is referred to many times in Kant's lectures on ethics, and it is discussed extensively in two major works from the 1790's (the *Religion Within the Boundaries of Reason Alone* and the *Metaphysics of Morals*), as well as in the essay on theodicy. As I have discussed previously, characteristic of these later works is their concern, not with the grounding of morality per se, but rather with the logically contingent but morally necessary features of our agency through which the moral principles can be applied. Significant roles are also

played by religion and moral education in the cultivation of moral dispositions as I shall explicate in the following chapters.

Taken collectively, these texts can be construed as parts of moral anthropology, which is of utmost importance if we are to make use of the moral principles presented in Kant's earlier works. That Kant was aware of the systematic importance of a moral anthropology is evident from the fact that almost all of his major writings from the 1790s are somehow related to it or can be used to expand upon it. We also see that conscience is discussed in detail in these texts.

Let us first look at the lecture notes on ethics in order to see the basic aspects of Kant's conception of conscience.

3.1 Conscience in the *Vigilantius Lecture Notes* (1793-4)

Before getting into the discussion on conscience, a short comment on the reliability of the lecture notes is necessary. The lecture notes taken by Kant's students and colleagues are usually not taken to be authoritative by themselves. The main reason for this is that Kant himself did not have his own manuscript for these lectures, and the notes seem to have been written after class. Hence it is likely that there have been some omissions from or distortions with regard to what Kant actually taught in class.[19] That being said, what we read from the lecture notes on the topic of conscience is exactly in line with what we can find in Kant's published works. For this reason, I believe the lecture notes can be used to ascertain some basic features of Kant's conception of conscience.

[19] See Denis & Sensen 2015, p.1-12 for further details on the reliability of the lecture notes.

Kant's first attempt to elucidate the concept of conscience can be found in the lecture notes taken by Georg Ludwig Collins in 1784. In these notes we find that Kant had already begun thinking about conscience as an internal court (CL 27:355)-an idea that still remains in the *Metaphysics of Morals* of 1797. However, his mature view of conscience is not to be found in these earlier notes, since he then thought that there could be errors involved in conscience. In his expositions of conscience during the 1790s he consistently defends the (highly unorthodox) view that conscience is infallible. For this reason I begin my discussion with the Vigilantius lecture notes.

The lecture notes taken by Johann Friedrich Vigilantius begin nine years later than the Collins notes, on October 1793, and consist of Kant's presentation of the metaphysics of morals. However, taken as a whole there are significant differences between the lecture notes and the eventual work that Kant published in 1797.[20] One reason for this is that Kant taught a separate course on the doctrine of right, and hence the Vigilantius notes do not have an extended discussion of issues related to right. With regards to Kant's conception of conscience, however, we do not come across many significant changes, although there are differences in emphasis and the method of explication.

In these lecture notes we also find the famous court metaphor of conscience, however, since it is discussed more in detail in the *Metaphysics of Morals*, I shall take up that issue in the next section. The main role of conscience in these notes is the examination of "inner actions" of a moral agent, since they can not be "known to an

[20] See the introduction to the Lectures on Ethics by J.B. Schneewind from the Cambridge edition in 1997 (page xviii).

external judge" (VL 27:572). These inner actions include our disposition toward the requirements of morality. The phenomenon of conscience involves a self-reflexive judgment on the "morality" of our actions, rather than their "legality". This is a key feature of Kant's conception of conscience and it explains in part why Kant discusses this concept within the purview of the *Doctrine of Virtue* (instead of the *Doctrine of Right*) in 1797.[21]

In Kant's view, this internal court can be regarded also as an external one if a person believes in God and "accepts [him] as a judge" (VL 27:574). This is a peculiar way of putting the issue, but I shall clarify it when I return to the court metaphor in the next chapter, where the presentation of this metaphor is more comprehensive. There is an important point to be noted in this passage as it is relevant for Kant's account of conscience in general: the internal forum of conscience can not settle issues about human justice, since this is in the purview of the faculty of understanding and its determining (VL 27:574). One central feature of Kant's account of conscience is that it is closely related, yet distinct from, the power of judgment. I shall explicate the significance of this point especially in the last section of this chapter where I discuss Kant's insistence on the infallibility of conscience.

In the same passage we also find a crucial insight that is of utmost importance for Kant: even if a person does not believe in God, s/he can still have a conscience "in case s/he possesses moral principles as such" (VL 27:574). This view of conscience can be called a "cosmopolitan view", since it disregards differences in religious convictions and

[21] Nevertheless, it will turn out that conscience is related to the Doctrine of Right as well, since it is related to the imputation of actions.

instead places conscience primarily into the domain of morals. The relation between conscience and religion remains relevant, but it is now endowed with a particularly moral import. I shall demonstrate that Kant held this cosmopolitan view of conscience throughout his various expositions of this concept.

In these lecture notes, Kant defines conscience as "the ability to impute one's own *factum* (deed) to oneself" and conscientiousness as the "readiness to do this [imputation]" (VL 27:575). Conscience, in this regard, is understood as presupposing an objective obligation and is relegated to the role of strengthening the disposition to fulfill that obligation. This is also one of the passages in which Kant ties the violation of conscience with the "loss of one's entire moral worth" (VL 27:575). This point is important for understanding the relationship between conscience and personal integrity, an issue to be discussed in the fourth chapter.

Generally speaking, in a given moral dilemma, any choice that goes against our conscience would, in that instant, result in a kind of "self-denial" or a threat to personal integrity. I believe Kant does have something along the same lines in his mind when he insists on the infallibility of conscience. One can not be mistaken that the call of conscience is sincerely one's response as a moral agent thus constituted. It is a normative response, in the sense that it is a response that one regards as appropriate and necessary. However, the response of conscience is not yet sufficient for initiating action. This requires forming a maxim to act and hence volition.

How the imputation of a deed is connected to conscience can be understood from some remarks in later passages. One important point related to this issue is that

conscience is similar to apperception, involving the "consciousness of my will, my disposition to do right . . . consciousness of what duty is" (VL 27:614). What is characteristic of conscience, then, is that it contains an awareness of the content of my volitions (and hence maxims). In order to impute an action to myself I need to be able to regard myself as responsible for the action that I have done. This complements the ancient conception of conscience as *suneidesis* (sharing knowledge with oneself of a wrong done). Also, this is one of the passages where the link between conscience and reflexive judgment becomes apparent, since it is only through the use of reflexive judgments that I can conceive of myself as having acted intentionally.

Generally speaking, intentional action is the kind of action for which I am morally responsible. We also see the same point made in the *Metaphysics of Morals*. There, Kant implies that conscience is the condition of all duties as such (MS 6:407), since without it no one "would neither impute anything to himself as conforming with duty nor reproach himself with anything as contrary to duty" (MS 6:401). Imputation results from our own awareness that the action that springs from us is intentional, and hence we are morally responsible for its consequences.

The close connection that Kant draws between conscience and judgment should not lead us to the view that they are identical.[22] In the next page of the lecture notes, Kant criticizes Baumgarten for equating conscience with judgment as "subsumption of our doings under the law" (VL 27:616), a view also held by Thomas Aquinas (see *Summa Theologiae* Ia, q. 79, a. 13). Once again, Kant distinguishes clearly between that

[22] This is a view commonly found in secondary literature as I shall demonstrate in the last section of this chapter.

which falls under the purview of the understanding (which concerns the determination of what our duty is in a given situation) and our subjective awareness of whether we have in fact done what is our duty or not (and whether we have done it for the right reasons or not).

A related issue is the kind of temporality involved in the different roles that conscience can play with regards to action. In this connection,, Kant proposes a distinction between an examining and a judging conscience (VL 27:615). The former relates to present and future action while the latter relates to our past actions.

Examining conscience is related to a deed: we reflexively examine ourselves to appraise whether we have considered all the relevant information that pertains the situation at hand. In that respect it is necessary "to ensure that no object is present in the *factum*, and known to us, that has not been examined and taken into account (VL 27:614). Kant lists three dicta about examining conscience: self-examination, reaching subjective certainty that the examination is thorough, and being sincere in our judgment of ourselves and the situation at hand. This role of conscience is also clearly tied to the notion of reflexive judgment (VL 27:618). In its examining aspect, conscience relates to a situation in which we are about to act, that is, to present or future situations of moral action and this aspect of conscience has to be cultivated so that we can orient ourselves better in moral situations (VL 27:617). The reason for this is that conscience "reinforces our awareness" that we are in a "situation governed by laws of duty" (VL 27:619).

What follows from this is that conscience, in its examining aspect, is inherently reflexive, but we have to keep in mind that its influence on our will is not automatic. We

shall see this point clearly made in the *Metaphysics of Morals*. In that text, conscience is regarded as providing the consciousness of the verdict of reason (MS 6:438). This consciousness by itself doesn't ensure that we shall act on that verdict. In other words, conscience is the cursor by which reason influences our deliberation in a moral situation, but it never determines it completely, as it is up to us whether to incorporate the verdict of conscience into our maxim.

Judging conscience, by contrast, pertains to our past actions and this is where feelings of remorse or a "nagging conscience" come into play. In Kant's view, this nagging conscience can only be soothed by amending the wrong we have committed, rather than wallowing in self-anguish (VL 27:618). This is crucial even at the hour of death. We are obligated to do what we can to ameliorate any situation in which we may have done wrong. As Kant puts it: "even in death we must be meticulous in preventing evil consequences of our actions from arising after our demise, and so must not disdain even the seeking for forgiveness" (VL 27:619).

Kant reinforces this point by appealing to a religious metaphor, known as the "bridge of Serat". This is a bridge that all people must cross in the afterlife, but if they have not made a reckoning with their evil deeds, they can not enter paradise (VL 27:619).[23] I shall mention instances where Kant uses religious imagery, as it will turn out that, according to him, this is one of the ways in which moral principles can be made more visible or sensible, or brought closer to intuition (see Rel. 6:66).

[23] This conscientious worry at the hour of death is reminiscent of Plato's portrayal of the final words of Socrates in the Phaedo: "Crito, we owe a cock to Asclepius, make this offering to him and do not forget" (Phaedo 118a).

In the Vigilantius lecture notes, we also see Kant putting forward his contention that conscience is infallible, by distinguishing between an error of judgment and an "awareness of the wrongness of reasons" (VL 27:614). What is unconscientious is to "regard something as a right while knowingly holding it to be wrong". This means that we can be mistaken about the rightness or wrongness of the action, but we cannot be mistaken about whether we believe that action to be right or wrong. Here, the role of conscience is looking at our subjective disposition toward the deed in question, which includes consciousness of the fact that the appropriate kind of self-examination has taken place. We cannot be in error with regards to whether we have examined ourselves or not and it is primarily in this sense that Kant thinks conscience is infallible.

This point has a significant political and religious implication as Kant uses this aspect of conscience to argue against any and all kinds of religious persecution. As the argument goes, judgment can be in error with regards to religious convictions, but religion can not be a matter of complete certainty so no one can have legitimate grounds for persecution. Laws of positive (revealed) religion "can furnish no adequate grounds for deciding" the matter (VL 27:615). In short, since no one can possibly demonstrate and justify the ultimate truth with regards to God and the afterlife, it is always unjustified to blame people with being unconscientious in case they have a different religious conviction (or none at all). This is why it is significant to espouse a cosmopolitan view of conscience. It seems clear that Kant was preoccupied with this issue, as he discusses the same point in relation to conscience also in his essay on theodicy and his *Religion*.

The requirement for certainty in matters of conscience is reminiscent of the legal principle of "proving that a person is guilty of a crime beyond reasonable doubt". In Kant's view, if there is possibility of error, it is best to refrain from judgment. In making this point Kant quotes the famous dictum by Pliny: *si dubitas ne quid feceris non facias* (VL 27:615). He uses the same quote in Latin in the last part of his book on religion where he discusses conscience. I examine that text in the next chapter.

In this section I have presented some of the central aspects of Kant's conception of conscience. In his view, conscience is involved in the imputation of a deed to ourselves; it is distinct but closely related to the power of judgment, it is infallible and it ought to be cultivated, especially in its examining aspect. I now turn to Kant's discussion of conscience in the *Metaphysics of Morals*.

3.2 Conscience in the *Metaphysics of Morals* (1797)

Kant's conception of conscience is formulated and discussed most extensively in the *Doctrine of Virtue*, which is the second part of the *Metaphysics of Morals*. In the introduction of the *Doctrine of Virtue*, conscience is counted among the four "concepts of what is presupposed on the part of feeling by the mind's receptivity to concepts of duty as such" (MS 6:399). It is crucial to grasp what this formulation means to see what place Kant accords to conscience within our moral lives. The other three preconditions are moral feeling, love of one's neighbor and respect for oneself (self-esteem).

What is significant about these preconditions is that our consciousness of them "only follow from consciousness of a moral law, as the effect this has on the mind" (MS

6:399). We do not have a duty to acquire these preconditions, as they are constitutive of our form of moral agency. That being said, we do have a duty to cultivate and strengthen these preconditions so that they aid us in acting in accordance with duty.[24]

In his paper entitled "Moral Feelings in the Metaphysics of Morals" Paul Guyer (2010) translates the phrase "concepts of what is presupposed on the part of feeling" with "aesthetic preconditions" (p. 130fn). According to Guyer, there is a hierarchical relation between these four preconditions in terms of generality. Moral feeling is the most general as it is "what makes us susceptible to the general idea of acting in accordance with duty" (i*bid.* 141). The other three preconditions relate to increasingly specific aspects of our moral practice. Conscience is concerned with particular maxims (ibid. 144), love of one's neighbor and self-esteem "impel us to strengthen our natural disposition to sympathy" and can play a role "as proximate causes of particular actions" (ibid.150).

According to Guyer's account, the cultivation of the most general precondition (moral feeling), leads to the cultivation of the other three preconditions in order of specificity. It is unfortunate that Guyer uses causal locutions to explain the relation between these preconditions and action.[25] As I have discussed in the previous chapter:

[24] This point seems to create a tension for Kant's distinction between acting from inclination (*aus Neigung*) and acting from duty (*aus Pflicht*). The worry seems to be that, if any inclinations are involved in performing the right action, we can not have acted solely from duty. Paul Guyer resolves this tension by offering a different interpretation. According to him, the requirement of moral merit (which is acting solely from duty) has to be interpreted as the requirement that the agent ought "to do what is necessary in order to fulfil his duty" which may involve cultivating certain moral feelings that facilitate acting in accordance with duty. Acting from duty is then construed as a second-order intention to do what one can so that the requirements of morality are responded to. I believe this is a very plausible reading which can be reconciled with Kant's texts. See Guyer 1997, ch. 10, p. 380-1 for more details on this interpretation.

[25] See for instance (Guyer, 2010, p. 140): "moral feeling plays a causal role in the etiology of particular actions" or p136 "in the second Critique he recognizes only one causally efficacious moral feeling, the

when the initiation of action based on a maxim is considered, a causal account cannot provide the proper normative grounds of guidance, since they can only present us with what is the case.

The main concern for the aesthetic preconditions, therefore, cannot be their causal efficacy (or lack thereof), but rather their appropriate cultivation and strengthening with respect to our moral agency in general, that is, the proper attunement of these preconditions in accordance with the requirements of morality. In this endeavor we are to employ the reflexive aspect of judgment (as it relates to features of intentionality), rather than the determining aspect (generally related to causal or substantial theoretical judgments). The reason for this is that morally significant actions can only be recognized to be such on the assumption of freedom of action, which requires not a determining judgment about causal processes but a reflexive judgment about intentionality: that is, unless freedom is central to our self-understanding, we can not make sense of morality at all.[26]

Kant asserts that there can not be a duty to acquire these aesthetic preconditions, rather, all human beings possess them, simply in virtue of being semi-rational agents (MS 6:399). These preconditions can be regarded as constitutive of moral agency (as they have their source in our consciousness of the moral law) but it is not enough to possess them, one needs to be attentive towards them and cultivate them. This is another

---

feeling of respect". According to the Incorporation thesis, no feeling or incentive by itself can be efficacious in any sense, unless and until we incorporate them into our maxims (or take them as reasons for action).

[26] See the footnote in MS 6:379: "For we can explain what happens [in an action violating the moral law] only by deriving it from a cause in accordance with laws of nature, and in so doing we would not be thinking of choice as free. -But it is this self-constraint in opposite directions and its unavoidability that makes known the inexplicable property *of freedom* itself".

corollary of the incorporation thesis that I have presented in the previous chapter; no matter which feelings or preconditions are present in an agent, only through their incorporation into maxims can they play a role in moral action. What needs to be cultivated, then, are not only these preconditions, but also our responsiveness to them as free agents. As we shall see, the attunement of these preconditions can only be accomplished through moral education.

What is significant is that the effect of the moral law on our minds does not only have intellectual import, but also an emotional aspect. All four predispositions involve both the thought of duty and some kind of feeling. One reason for this is that the human will is "pathologically affected" but not determined (A534/B562). Another reason is that all our volitions (or indeed, all our experiences) involve some sort of inclination (G 4:398). As I have discussed in the previous chapter, moral feelings increase our awareness of the morally salient features in our everyday experience. This is the way in which Kant construes the aesthetic predispositions, and hence conscience.

I claim that the voice of conscience has to yield a particular affective tone through which it emulates a certain kind of moral experience. This is the way in which conscience can be understood as distinct from other mental phenomena. In order for us to attribute this experience to conscience, we need to be able to discern what distinguishes conscience from other moral feelings or mental phenomena. Kant associates the feeling of awe (respect coupled with fear) with the functioning of conscience (MS 6:438).

Conscience provides an affective link between the moral law and our minds-I recognize this law as authoritative, and therefore I recognize myself as obligated to act in accordance with it. This imputation has both an intellectual and an affective character. If conscience were purely intellectual, a further question would arise as to that which gives rise to the painful feelings of guilt and remorse. These feelings must be understood as part of the operation of conscience, or we would need a further capacity to translate the verdict of reason into such feelings apart from conscience. The talk of "pangs" of conscience also relates to this point (See VL 27:43). Otherwise, we face the danger of conscience being swallowed up by our intellectual capacity-reason.

What makes conscience distinctive, then, is that it incorporates both feeling and intellect. However, the same is true of what Kant described as the feeling of respect in the 1780s, and what he calls the four aesthetic preconditions in the *Metaphysics of Morals*. After the passage in which Kant discusses the court metaphor of conscience (a topic to be discussed shortly), we get a clue as to what distinguishes conscience from the other aesthetic preconditions: "Every human being has a conscience and finds himself observed, threatened, and, in general, kept in awe (respect coupled with fear) by an internal judge" (MS 6:438).

As we have seen, Kant usually correlates the general feeling of respect with our consciousness of the moral law. Conscience, then, adds an element of fear, which can be attributed to a fear of punishment. This fear of punishment can only arise in a situation where there is danger of deviating from the requirements of morality and hence it can be felt prior to the deed, perhaps via a representation of the deed in question. What we fear in this situation is to be found guilty of a violation of duty (MS 6:440). In this context,

what distinguishes conscience from the moral feeling is that it is more specific and it incorporates an element of fear.

Kant's remarks on the feeling of awe (*Ehrfurcht*) will be helpful to understand the kind of feeling that is associated with conscience. In his essay on the relation between theory and practice (dated 1793), Kant relates the feeling of awe with our recognition of the greatness and sublimity of our true vocation. This awe accompanies an inner experience in which "the mind is elevated and animated toward a pure moral disposition" (TP 8:287). Here Kant also suggests that, in private and public instruction, we must draw attention to the fact that we are able to do as duty requires with appropriate respect toward the moral law.

In the *Metaphysics of Morals*, the feeling of awe is associated with the respect that is owed to the "principle of God's right" which is "justice" (MS 6:489). Here, justice must not be construed as something that we owe to God, since, ultimately, we can only comprehend the moral relations of human beings to human beings (MS 6:491). Nevertheless, the pursuit of justice in this world has an affective aspect to it, which includes feelings of awe and reverence.

The feeling of awe, however, is not necessarily connected to a particular religious doctrine or denomination. This point is made clear by Kant in his book on *Religion*. As he puts it "[. . .] awe is not a particular act of religion, but the religious disposition which universally accompanies all our actions done in conformity to duty" (*Rel.* 6:154fn). In this work, Kant also describes the feeling of awe as being instilled by the majesty of law. This feeling rouses "the respect toward a master . . . that lies in us"

and hence it is ultimately a feeling directed at the "sublimity of our own vocation" (Rel. 6:23fn).

The feeling of awe can be shown to be associated with what Kant calls reflexive judgments on the sublime in the third *Critique*. While Kant does not spell out a direct connection between conscience and the sublime, I believe his remarks allow for an interpretation which ties them together. That being said, one crucial difference has to be noted: reflexive judgments on the sublime constitute an aesthetic judgment, which Kant categorizes as disinterested and hence not directly related to the faculty of desire: it is instead related to the enlargement of the faculty of imagination and its harmony with the faculty of reason (KU 5:250, 5:256). Conscience, on the other hand, is always involved with the practical aspect of reason and hence is "interested" (VL 27:620).

Nevertheless, the affective aspect of these operations of the mind seem to carry some similarity, as they both give rise to the feeling of awe and they are both related to the effects of the moral law upon our minds. The judgment of the sublime "awakens the feeling of a super-sensible faculty in us" (KU 5:250). The super-sensible here must be understood as an authority that rules over our sensible nature, since Kant associates the feeling of sublime as related with the "dominion that reason exercises over sensibility only in order to enlarge it in a way suitable for its own proper domain (the practical) [. . .]" (KU 5:265). Recognizing the authority of the moral law that arises from reason, then, gives rise to the feeling of awe. From an interested, that is, practical perspective this feeling is associated with conscience. From a disinterested, aesthetic point of view, this feeling is associated with reflexive judgments on the sublime.

In one of Kant's earlier lectures on ethics we find the relationship between awe, conscience and religion made clear. In the notes taken by Herder, awe is related to reverence, "together with an anxiety not to offend its object". This concept of awe gives rise to a specific kind of fear, "against an evil which we do ourselves" (HL 27:32) . This is clearly the anxiety associated with conscience, experienced in a situation in which there is a danger of violating what we take to be our duty.

In the Vigilantius lecture notes, Kant also relates the feeling of awe (which he also calls supreme respect) with piety (*pietas*), which is "the disposition to perform virtuous actions in a godfearing frame of mind, representing the highest stage and a pendant to duty, since human duties are here construed as commands of God" (VL 27:715). The feeling of awe is ultimately connected with what we take to be sacred. As I shall demonstrate in the fourth chapter, Kant's conception of conscience is intimately related with holiness, but construed not from the perspective of a revealed religion, but rather from the perspective of a more cosmopolitan, moral religion..

An issue related with holiness can be found in the *Metaphysics of Morals* as well. Conscience, according to Kant, appears peculiar because its dictates seem to be those of another person, although one and the same person is both judge and s/he who is judged in the internal court of conscience (MS 6: 438). This passage implies a duality of persons residing in one and the same subject.

Kant grounds this distinction in his Transcendental Idealism, claiming that the "judging" aspect is done by the "homo noumenon" while that one who is being judged is the "homo phenomenon" (MS 6:335, 6:418). However, he doesn't need transcendental

idealism to provide a basis for this distinction, he only needs a conception of moral agency that is capable of acting from duty (autonomous), but that is also prone to deviating from the moral law (heteronomy). This involves incorporating an awareness of our moral worth and comparing it with a moral ideal within the phenomenon of conscience. As I have shown in the previous chapter, ideals arise from our reflexive judgments pertaining to a systematic unity of our concepts, ends or principles. Kant already draws attention to this ideal in connection with conscience in the *Metaphysics of Morals* (MS 6:438). In order to see how this moral ideal functions, let us take a closer look at the court metaphor that Kant puts forward in the *Doctrine of Virtue*:

> Every concept of duty involves objective constraint through a law (a moral imperative limiting our freedom) and belongs to practical understanding, which provides a rule. But the internal imputation of a deed, as a case falling under a law (in meritum aut demeritum), belongs to the power of judgment (iudicium), which, as the subjective principle of imputing an action, judges with rightful force whether the action as a deed (an action coming under a law) has occurred or not. Upon it follows the conclusion of reason (the verdict), that is, the connecting of the rightful result with the action (condemnation or acquittal). All of this takes place before a tribunal (coram iudicio), which, as a moral person giving effect to the law, is called a court (forum). Consciousness of an internal court in man ("before which his thoughts accuse or excuse one another") is conscience. (MS 6:438)

This passage reinforces Kant's distinction between the process of judging and our consciousness of it. During moral deliberation, there are various thoughts at play, representing different points of view on the matter at hand. At the end of this inner discussion, reason reaches a verdict and this verdict is made known to us through a verdict, which either carries with it an affective probing or "nagging" or relieves us from our anxiety of being found guilty in our own regard. That conscience is intimately

connected with sensibility is apparent. As Kant puts it: "although the pain one feels from the pangs of conscience has a moral source it is still a natural effect, like grief, fear, or any other state of suffering" (MS 6:394).

Now, in our eventual action we may choose to listen to our conscience, or we may choose to do something else: this would result in following one of the thoughts that "made its case" in the court of reason. In case this happens, we are not acting sincerely; that is, we are not acting according to our best judgment. As we have seen in the Vigilantius lecture notes, this is "regarding an action as right while knowing that it is wrong" and hence it is blatant insincerity (VL 27:614).

In his discussion of court metaphor, Kant appeals to God, or a subjective representation of God as the inner peer mentioned earlier (MS 6:439). This inner peer is the representative of what we take to be a moral ideal. We ought to compare our actions and maxims with what we think would be the actions and maxims of a morally perfect being, imagined as acting always from duty (*aus Pflicht*) and never from inclination (*aus Neigung*). Kant insists that we must not compare ourselves with other people in terms of moral perfection, but we must rather compare ourselves with the requirements of morality and to what extent we have been successful in meeting them (see CL 27:349; KpV 5:37, 5:69, 5:74; MS 6:435, 6:437).

In his explication of Kant's account of conscience, Owen Ware suggests that what the agent takes to be the moral ideal in conscience is "who s/he ought to be" (Ware 2009, 691). This sounds plausible, but I believe Kant is searching for a stronger ideal. What judges us in conscience is not only our best version, but also a "scrutinizer of

hearts" (MS 6:439). Kant is usually pessimistic with regards to self-knowledge (MS 6:382, 6:447). This means that we need to imagine an agent who has a better access to our inner dispositions than we do. This could only be conceived as an ideal spectator, or God.

Even though we can not know our deepest dispositions with certainty, we can strive to make them accessible and also try to model our behavior on what an ideal agent would do in a given situation. We can not know our own maxims, since we are prone to self-deception (Rel. 6:68, T 8:268fn), but we can know what an appropriate and moral maxim would be, since we are aware of the requirements of morality. This moral ideal can be used as one precisely because it can not be compared with anything else, since it is of supreme significance and hence is incomparable and inviolable (G 4:435). We can not measure a measuring-rod without reference to that measuring-rod. In this sense, the moral law is the sole principle we can use for a precise measurement of legitimate action and moral disposition.

In conscience, we imagine this being also as judging us, as it knows all of our dispositions (MS 6:439). Here, conscience can be understood as affective feedback through which we assess our own judgments-we imagine a morally superior person (a judge) who has full access to our motives and maxims and we submit our purported (or past) decisions to its scrutiny. This judge is a figure which we ought to respect and fear. This is why Kant characterizes conscience as "being accountable to God for all of one's deeds" (MS 6:439). This accountability to God is closely related with Kant's definition of moral religion: "the recognition of all our duties as divine commands". This is an issue I take up in the next chapter.

Before getting into the relation between religion and conscience, I shall discuss one of the most central issues in the interpretation of Kant's account of conscience, namely his insistence that "an erring conscience is an absurdity" (MS 6:401). After presenting my own interpretation, I shall consider and criticize some secondary literature that pertains specifically to the infallibility of conscience. In this way I shall demonstrate the merits of my own interpretation of Kant's account of conscience.

3.3 Infallibility of conscience

Kant makes two important claims about conscience which have to be discussed in some detail. Conscience "speaks involuntarily" and it "can not err" (MS 6:401).[27] As I have argued, these properties seem to put conscience more on the side of receptivity than spontaneity. That being said, we need some control over our conscience so that it is possible to cultivate it. While an involuntary and un-erring response can hardly be said to be related to anything intellectual or spontaneous, the ability to cultivate seems to imply a middle-ground between receptivity and spontaneity. I shall attempt to show that it is in this middle-ground where the infallibility of conscience lies.

In Kant's view, spontaneity is the common property of reason, understanding and judgment and these faculties are clearly distinguished from sensibility, which is generally construed as receptive (B68, A51/B75, B93, A97). In this respect, our

[27] As Thomas Hill (2002) puts it: Conscience is an involuntary response to the recognition that what we have done, are doing, or are about to do is contrary to the moral judgments that we have made (by applying moral law to different types of circumstances) (p. 303)

judgments are fallible, but feeling is infallible (A293/B350).[28] As a corollary, we could state that the spontaneous aspect of cognition is always open to mistakes. Hence, one reason for the infallibility of conscience could be that we do not have spontaneous control over it.

Some commentators have suggested that, since Kant construes conscience as infallible, it can not play a central role in his moral philosophy. Andrea Esser (2013)stresses that "[i]nsofar as this characterization of conscience introduces an element of subjective immediacy into Kant's critical ethics, the conception of conscience remains problematic." (p. 277). Given Kant's synthetic view of cognition and his insistence on the fallibility of judgment, it must be admitted that the prospect for an immediacy involved in conscience does not seem to fit well with Kant's understanding of the spontaneity that has to be involved in moral thought and action. I believe this problem can be remedied if one can show in what sense conscience is infallible, thereby properly placing this special kind of immediacy within Kant's framework.

Infallibility is perhaps the most extensively discussed issue about Kant's conception of conscience. Asserting that conscience is infallible contradicts many views on conscience that came before (and after) Kant. What is more, we also come across the view, in the *Metaphysics of Morals*, that a person who acts according to his/her conscience, can not be guilty. As Kant writes:

> But if someone is aware that he has acted in accordance with his conscience, then as far as guilt or innocence is concerned nothing more

[28] See also CF 7:100 and AP 7:146: "Senses do not deceive. Error is a burden only to the understanding".

> can be required of him. It is incumbent upon him only to enlighten his understanding in the matter of what is or is not duty; but when it comes, or has come, to a deed, conscience speaks involuntarily and unavoidably. (MS 6:401)

Let us try to see why Kant thought this way and how he tries to make this point. As we have seen, conscience is the "consciousness" of the inner court and hence cannot be equated with judgement itself (MS 6:438). The function of conscience, then, is to raise in us the awareness of: a) the imputation of a deed to ourselves which involves b) relaying to us the verdict of reason in a specific situation. Since we have cognize the requirements of morality, presented as verdicts of reason in conscience, we are morally responsible for complying with them in our intentional actions. Taken in this sense, it seems plausible that conscience, as a faculty of awareness akin to perception, can not be in error. In other words, conscience is related to an awareness of our moral beliefs and standards that bear upon the specific situation at hand.

This does not mean we are not responsible for our "honest mistakes", it merely means that acting conscientiously is acting according to our best judgment. Our best judgment may still be mistaken, due to a variety of contingencies, but we can hardly be mistaken that it was indeed our best judgment. Our best judgments, in turn, have to incorporate reflexive judgments as to the adequacy and extent of our moral self-examination. In the Theodicy essay we find the distinction between the judgment of understanding presented in contrast with the "voice" of conscience. Kant states that we could be mistaken with respect to the former but not the latter. As he puts it:

> For in the first instance (the truth or falsity of a statement) we compare what we say with the object in a logical judgment (through the

> understanding), whereas in the second instance, where we declare what we hold as true, we compare what we say with the subject (before conscience). (T 8:267)

What comes to mind is something similar to the basic features of perception: I can be mistaken in whether or not the object I see before me is actually there (due to some optical illusion or a malfunction of the senses), but I can hardly be mistaken in whether it seems that the object is there. The subjective experience of conscience, then, is not open to doubt.

I cannot be mistaken in what I hear as the voice of conscience itself. Whatever may be the pronunciation of the inner court, I can not be mistaken that it is in fact the verdict of that court. As stated in the previous section, I believe it is in this sense that Kant claims an "erring conscience is an absurdity" (MS 6:401). This can only be explained if the voice of conscience is "distinctive" and "unique", its call must be distinguishable from any other thoughts that we may happen to entertain. As I have shown, this uniqueness can be attributed to the feeling of awe defined as "respect coupled with fear", specifically functioning alongside the operations of conscience. We especially become aware of this affective aspect in the acquittal or condemnation of conscience.[29]

If I listen to my conscience I can not be held guilty of dishonesty, in other words, acting according to conscience is acting with the best intention I have and the best

[29] In his paper on Kant's account of conscience, H.J. Paton (1960) writes: "It (conscience) asks whether we honestly believed that our action was not wrong and whether we took the trouble necessary to justify our belief. If the answer to these questions is an honest affirmative, we are acquitted. If it is negative or even doubtful, we are condemned" (p. 242).

judgment I have reached. I think this is what Kant has in mind when he says that no more can be required of a person who has listened to her conscience. This does not imply that the verdicts we have reached through rational self-deliberation can not be mistaken at all. In fact, Kant always warns us about the possibility of self-deceit, as well as the opaqueness with regard to ourselves (TP 8:284 and MS 6:446-7).

Our best judgments will be the ones which take all the relevant factors of a given situation into consideration. It will involve thinking without relying on an external authority, but also thinking from the standpoint of others and doing so consistently. Finally, our conscience will relay to us the final verdict of reason and we will have acted to the best of our abilities–which never guarantees success but only sincerity, which Kant thinks is the best we can hope for.

Throughout his writings on conscience, Kant insists that we are immediately aware of what we hold to be true. We owe this immediate awareness to conscience, since conscience "reports" to us that the judicial process of reason has reached a verdict through a sufficiently comprehensive examination. Through the operation of conscience, reason reaches a verdict about itself; what is at stake here is not primarily the truth or falsity, but rather the sincerity of our judgment. In other words, our conscience resonates with what we sincerely hold to be true. This is what it speaks. In case we say anything different than what our conscience repeats back to us, we are in effect lying, since we must be saying something which we do not hold to be true. As we have seen in the Vigilantius lecture notes, this amounts to unconscientiousness. This results in a

disharmony between how we think and how we act (or speak).[30] Kant presents the same point thus in the Theodicy essay:

> One cannot always stand by the truth of what one says to oneself or to another (for one can be mistaken); however, one can and must stand by the truthfulness of one's declaration or confession, because one has immediate consciousness of this. (T 8:267)

This is one of the instances in which our conscience may reproach us. We can read an interesting discussion about the nature of the reproach of conscience here. Kant contends that "every crime already carries with it its due punishment, inasmuch as the conscience of the perpetrator torments...even more harshly than the Furies" (T 8:261). We have seen this aspect of conscience also in the Vigilantius lecture notes where Kant spoke of a "nagging conscience". This idea is reminiscent of Socrates' argument in the *Crito* about how "wrongdoing and injustice is in every way harmful and shameful to the wrongdoer" (*Crito*, 49b). When we do something morally wrong, we are in effect harming ourselves as well-if our conscience is functioning properly.

This point, however, can not be pursued too far, since the person committing the crime might be so depraved as to have completely shut off the voice of his/her conscience. Kant thinks that conscience is so pervasive that we can never silence it. As he puts it: "it follows him like his shadow when he plans to escape[...] He can at most, in extreme depravity, bring himself to heed it no longer, but he still cannot help hearing it." (MS 6:438)

[30] As Kant puts it: "by making a declaration of whose truth he is not convinced, something contrary to his conscience" ® 6:171).

Kant also warns us against imputing the disposition of a virtuous person upon someone who has become excessively evil. This shows that conscience is not some a priori capacity which functions somewhat adequately in all human beings; rather, the effectiveness of conscience is contingent upon the moral character of the person in question, and upon whether she incorporates the dictates of her conscience into her maxims. This requires that one cultivates one's conscience in order to "sharpen one's attentiveness to the voice of the inner judge and to use every means to obtain a hearing for it" (MS 6:401). On the other hand, for the person who chooses not to cultivate her conscience in this way, "the scrupulousness of the honest who inwardly plague themselves with self-inflicted rebukes" is merely laughable (T 8:261).

Generally speaking, the less conscientious a person is, the more s/he is invested in a short-term satisfaction of desires as that kinds of person will typically act from inclination rather than from duty. To the extent that we stop listening to our conscience, which presents the verdict of reason, we are bound to dull our rational/moral capacities, as we stop listening to the demands of morality and rather become prone to adopting a principle of self-love, which underlies our individual maxims. This point will be discussed under the heading of radical evil in the next chapter.

Kant's warnings about listening to the dictates of conscience emphasizes the necessity of this precondition in order for anyone to become a moral being that is capable of acting resolutely and effectively. As we have seen, conscience has the function of imputing a deed to ourselves and hence holding us morally responsible for our actions. However, as the aforementioned passage makes it clear, acting in accordance with the voice of conscience is a further issue. Again, after we hear the voice

of conscience, it is up to us to incorporate our best judgement into our maxims, which then leads to intentional action.

In order to elaborate this point further, and to tie the discussion of maxims with the discussion of infallibility, we may think of a distinction between the voice of conscience and its content. Since the content carried with the voice of conscience consists in the outcome of the proceedings of the inner court, it necessarily involves a judgement of the understanding and hence is fallible. However, the voice of conscience, since it is a direct cognitive and emotive response of our moral constitution, speaks directly and infallibly.

As we have seen, our moral character involves our core convictions about morality, or what we take to be true in the moral realm. In his paper on Kant and conscience, Claudio La Rocca (2016) explains the kind of conviction that is at stake in conscience by referring to a passage in the first *Critique* (p. 73). In this passage, Kant is making a distinction between two senses of holding something to be true; if it is valid for everyone, the belief is called conviction, but if it "has its ground only in the particular constitution of the subject, then it is called persuasion" (A820/B848).

I contend that the voice of conscience is a voice of persuasion that "sounds like" a conviction and this is the sense in which we cannot be mistaken. This point might be confusing, since according to Kant, we can not subjectively distinguish between conviction and persuasion. Furthermore, this idea can come into conflict with Kant's insistence that conscience can not err, since we are capable of a kind of self-deception in which we can "feign conviction" even in our "inner professions" (T 8:268fn).

Indeed, Kant always stresses that we are prone to deceiving ourselves, or being too lenient in our self-appraisals. This leniency stems from our ineliminable self-love and it is why Kant thinks that human beings are evil by nature ® 6:70). However, we can also feign conviction in order to escape punishment. In the *Religion*, Kant discusses why it is absurd to force a confession before a civil tribunal (6:159fn), since one's convictions are the business of conscience, forcing someone to utter something that one is not convinced of can only lead to hypocrisy.

A persuasion that sounds like a conviction seems puzzling in itself. However, we could draw an analogy to the effect that this kind of persuasion is akin to a judgment of taste; it has a universal import, even though its ground is in the particular constitution of the subject. The distinguishing mark of a judgment of beauty is that it relays a specific kind of pleasure that arises from the harmonious free play of our imagination and understanding (KU 5:218). The analogy here would be a judgment that is relayed to us by conscience involving another kind of feeling: the feeling of awe. Without some distinguishing feature (or perhaps a specific phenomenology) Kant would have difficulties in asserting that conscience is infallible, since we can not differentiate between conviction and persuasion subjectively.

What we can know with certainty is what we regard to be right and wrong. Since conscience concerns sincerity, the cultivation of conscience must in part consist in the cultivation of sincerity. Sincerity requires that we are always disposed to asserting what we hold to be true. This point has implications for the method of moral education that is appropriate to the cultivation of conscience. I shall take up this issue in the fourth chapter.

3.4 Secondary literature on the infallibility of conscience

Let me now consider some arguments found in the secondary literature in order to ascertain whether my interpretation of the infallibility of conscience is plausible. Contrary to many commentators, I believe this infallibility is rooted in conscience being intertwined with sensibility. In the secondary literature, the main strategy has been to interpret this infallibility as being due to a special kind of judgment, a reflexive judgment in which the agent makes sure that she has been diligent in her judging. As I have demonstrated, while it is true that conscience is intimately related with reflexive judgment, they must be kept separate. Kant contends that the diligence involved in making any kind of judgment has to be present in any and all kinds of "rational" judgment (recall that rational judgment is structured normatively and self-reflexively as I have discussed in the previous chapter). Conscience can not be identified with a special kind of moral judgment (as Knappik& Mayr, Kazim and Vujosevic contend) and therefore it can not account for the infallibility of conscience.[31]

For these reasons, the infallibility of conscience has to be accounted for not in purely intellectual terms but with terms that also appeal to sensibility. The faculty of judgment is always fallible and, as we have seen, Kant is careful in distinguishing conscience from it. We hear the voice of conscience unavoidably and involuntarily (MS 6:401). Furthermore, we have an affective experience with regard to conscience–its pangs, its threats of punishment, its keeping us in awe etc., but we need not have any affective aspect in our usual cognitive functioning.

[31] At this point it can be helpful to return to points I made earlier about the normative character of rational judgment.

Dean Moyar (2008) argues that Kant's conception of conscience creates a tension for his moral theory in general. This tension concerns the primacy of conscience in Kant's account. In the first instance, conscience is a prerequisite for being a morally responsible agent in the first place. On the other hand, since conscience presupposes moral judgment (among other features of practical reason), it can not function without it (Moyar 2008, 341).

As we have seen, Kant claims that conscience is an integral part of our constitution, it is a natural part of human agency, it enables us to hear the decisions of the inner court (which is practical reason). However, Kant does not say that conscience is the only thing constitutive of our moral agency, nor does he say that it has priority over everything else that pertains to moral agency.

Moral judgment itself presupposes some knowledge of morally relevant features in specific situations (what Barbara Herman (1993, p. 77) calls rules of moral salience). What this means is that making any moral judgment, and, in turn, having a well-functioning conscience depends on our previous moral education, our awareness of a world with moral features. Moral judgments, imperatives, duties and conscience; all of them come into view after we have received some sort of moral upbringing. We do not start engaging in morality from scratch.

Another reply to this point can be given with the distinction I made earlier about the *ratio cognoscendi* and *ratio essendi*. One of the ways in which we experience the effects of the moral law upon our minds is conscience, since it is one of the four aesthetic preconditions of the mind's receptivity to moral law. This means that

conscience has no role in assessing the legitimacy of the moral law (or that which morality requires), but rather it "prepares" the person to be responsive to the demands of morality. This preparation also involves prior education and maturation, not only of our intellectual capacities but also of our sensible nature. When we put the picture in this way there is no circularity about the systematic role that conscience plays–it is essential for becoming a fully-functioning moral agent.

Broadly speaking, there are two main lines of interpretation in the secondary literature on Kant's conception of conscience. Some commentators take conscience to be "intellectual" and as "having an effect on feeling". In her paper on Kant's account of conscience, Marijana Vujosevic (2014) argues that conscience is a specific manifestation of practical reason. As such, it is a "kind of moral self-assessment that involves cognizing and judging our own character" (p. 450). Ultimately she claims that conscience is only intellectual as it can not be mere feeling.

There is some textual evidence for this point; in the *Religion* Kant says that "should anyone fear that his reason, through conscience, will judge him too leniently, he errs, I believe, seriously" ® 6:70fn). Also, in the *Metaphysics of Morals*, conscience is characterized as "an original intellectual and moral predisposition" (MS 6:438). It is obvious that conscience can not merely consist of feelings. If that were the case, we would not intelligibly talk about cultivating conscience, just as we can not make sense of cultivating pain.

That being said, as I have demonstrated, Kant consistently distinguishes between the operations of reason, understanding and judgment from the functioning of

conscience. Arguing that conscience is purely intellectual is inconsistent with Kant's repeated claim that it is infallible. Even in the passage above from the *Religion*, it seems that the judging is done by reason alone, and conscience is a kind of "tool" of reason to exact a kind of "torment or reproach" to the person. In order for conscience to fulfill this function, I believe that it should incorporate elements of both an intellectual and an affective nature. Indeed, this is the mark of all aesthetic predispositions discussed in the *Metaphysics of Morals*, since they consist of a receptivity to concepts of morality.

Vujosevic (2014) points to a passage in the *Lectures on Ethics* in which Kant says that conscience 'conveys an inner pain at evil actions, and an inner joy at good ones' (CL 27: 297). Conscience, according to her, operates by "linking a particular feeling with the action" (p. 457). There are two problems with this view. The first problem is that this quotation is taken from the Collins lecture notes, in which Kant had not yet developed his mature view of conscience (as I have shown in the first section of this chapter).

The second problem is to explain how conscience could "link a feeling with the action", if it had no relation to the affective side of our constitution. The second camp in the interpretation of Kant's account of conscience can answer this problem, since they generally argue that conscience is a kind of feeling. According to Allen Wood (2008), "conscience is a feeling of pleasure or displeasure associated with myself, in view of some action I am either contemplating or that I have already performed" (p. 183). This line of thought, I believe, is also mistaken, as it would be difficult to reconcile the mere feelings of pleasure and displeasure with issues about sincerity and imputation.

Knappik and Mayr's (2013) discussion of conscience also seems to put them alongside Vujosevic's interpretation, as they rely heavily on an account of proper moral judgment, which they argue requires diligence and certainty. These are traits that conscience is supposed to inspect. They claim that certainty should be regarded in a "non-factive" way, not requiring truth. This much is certainly correct, since conscience involves subjective certainty. They also provide a detailed and plausible account of moral judgment, but they fail to show how judgment is connected to conscience. This is quite striking, as they begin their paper by showing how Kant himself distinguishes between conscience and judgment. To disregard this claim and assert that conscience "issues an infallible second-order judgment whose object is the first order moral judgment of understanding" is to miss the point about infallibility (p. 15).

In his book on Kant's account of conscience Emre Kazim also defends the view the conscience is primarily intellectual. Kazim (2017) cites the definition of thinking from the first *Critique* (A69/B94) and claims that conscience is also a type of thinking. The definition reads: "Thinking is cognition through concepts". He then distinguishes between two aspects of conscience: intellectual conscience, defined as "the judgment of the internal court", and it's effect on moral feeling as the "consciousness of this judgment which motivates the agent". This interpretation also faces the difficulty of explaining how, on Kant's understanding of our cognitive faculties, there could be "involuntary" or "unerring thinking".

As I have shown, the correct way to account for the infallibility of conscience is not to endow it with a special kind of judgment. Any and all judgments issue from the same faculty of judgment. Conscience is infallible not because it engages in a second

order judgment; it is infallible because we can not use it spontaneously as we do the faculty of judgment. Conscience is best understood as the affective way in which we register the normative force of moral considerations by comparing our inner dispositions with that of an ideal moral agent. The feelings associated specifically with conscience, such as awe, reverence and fear of doing wrong, can be cultivated or ignored altogether but we can not give rise to them by sheer force of will.

For these reasons, any account which identifies conscience with judgment can not make sense of Kant's assertions that conscience speaks "involuntarily" and "unavoidably" (6:401). An involuntary and unavoidable judgment would be a chimera, at odds with Kant's claims about spontaneity and freedom with regard to the faculties of the mind. An account which takes conscience to be merely a kind of feeling, on the other hand, would have problems in explaining its relation to imputation and sincerity.

The interpretations of Kant's account of conscience found in the secondary literature has not tied it specifically to the feeling of awe or to reflexive judgments. Furthermore, I believe my interpretation is the first to connect Kant's insistence on the infallibility of conscience with his wider political aims regarding cosmopolitanism. In the next chapter I return to this issue as I examine Kant's understanding of the relationship between conscience and religion.

In this chapter I have focused on two texts in order to ascertain Kant's general conception of conscience as it relates to our moral and cognitive capacities. I began with a discussion of the lecture notes taken by Vigilantius, where some of the key aspects of Kant's conception of conscience could be found: we have seen that conscience is a)

concerned with imputation and our inner dispositions, b) that it is infallible, c) that it is distinct from yet related to judgment and d) that Kant espoused a cosmopolitan view of conscience.

In the next section I examined Kant's conception of conscience in the *Metaphysics of Morals*. I began this discussion with what Kant calls aesthetic predispositions. Crucially, these are moral feelings which incorporate both the thought of duty and some form of feeling. I argued that what distinguished conscience from the other predispositions was its unique connection with reflexive judgment and its ability to give rise to a feeling of awe-a point which can be tied to the discussion in the third *Critique* regarding reflexive judgments on the sublime. I have also discussed what Kant means by the infallibility of conscience against the background of some discussions in the secondary literature and have shown how my interpretation has significant advantages over some other attempts of exegesis. In the next two chapters I shall turn to the more specific aspects of conscience, namely its relation with religion and its role in moral education.

CHAPTER 4

CONSCIENCE AND RELIGION

In this chapter I shall discuss two texts that Kant wrote during the 1790's, namely *Religion within the Boundaries of Pure Reaso*n (1793) and his essay entitled "On The Miscarriage of All Philosophical Trials in Theodicy" (1791). In these texts Kant is mainly concerned with principles of biblical exegesis. However, he does have a lot to say about morality, human nature and conscience, as these are all bound up together in his discussion of religion. Kant is certainly not of the opinion that religion is purely a matter of faith or conviction, but rather he takes it to be a primarily moral endeavor. For this reason, according to Kant, biblical exegesis must also follow suit and lend its resources to the moralization of humanity (Rel. 6:112).

4.1 Conscience in the *Theodicy* essay (1791)

The text on *Theodicy* is mainly about the principles that Kant thinks ought to be employed in biblical exegesis. Some relevant aspects of this text have already been discussed in the previous chapter, such as the distinction Kant draws between the voice of conscience and its content. We can also find remarks concerning sincerity and conscience in this text. What is important for my purposes in this chapter rather concerns the relation between conscience and religion.

Religion is a matter of conscience in the sense that everyone has to decide for themselves which religious doctrine one may commit oneself to-or not commit at all. On

the one hand, it is immature for someone to simply adopt a given religion without reflecting on it (WE 8:35 and also NF 18:602). On the other hand, anyone who puts forward some assertoric statement about religious doctrine (which, if institutionalized, could also lead to the persecution of those who do not hold the same belief) is trying to go beyond the boundaries of reason ® 6:154fn, see also O 8:138fn).

As we have seen, in the Vigilantius lecture notes Kant made the case that all religious persecution is impermissible, on the grounds that positive (that is, historical/revealed) religion can not provide the grounds for deciding the matter of the truth or falsity of religious beliefs (VL, 27:614. See also Rel. 6:154). This is one of the reasons why engaging in a critique of reason ought to be not only a theoretical, but also a "political" activity (see O'Neill, 1989, 4). A political or religious institution which purports to assert objective certainty about a religious doctrine could thus be opposed by appealing to the critique of reason. As I have shown in the previous chapter, Kant argues that issues regarding the truth of religious convictions can not be settled rationally or justifiably, which is why all religious persecution is impermissible.

In the text on theodicy, Kant engages in a discussion of the *Book of Job*. This discussion is fruitful in understanding the link between religion and conscience, as well as the interpretative strategies that Kant wishes to employ in biblical exegesis. As we shall see, these points are further elaborated in his book on religion.

A peculiarity should be noted here first. Kant's commentary on the *Book of Job* seems to indicate that he finds, in the character of Job, an exemplar of conscientiousness, honesty and true/natural religious disposition (T 8:266). In the book on *Religion* (written

two years after this text) Kant elaborates his account of true or natural religious disposition and claims that Christianity is the closest faith to anything that could be called "true religion" in his sense of the word (Rel. 6:84). By contrast, his comments about Judaism in the *Religion* are anything but laudatory; claiming that it relies too heavily on external observances and that the Israelites had a merely "political" connection to their God (Rel. 6:126). For this reason, it is strange that in his essay on theodicy we find that a book from the Hebrew Bible is presented by Kant as an example of conscientiousness.

Kant stipulates that the most important aspect of the "completeness" of Job (apart from the goods and riches that he enjoys before the tribulations he faces) is that he is at peace with himself in a "good conscience" (T 8:265). In this context, this means that he has never sinned against God. If we consider the court metaphor from the previous chapter, never sinning against God amounts to a alignment between Job's moral ideals and his actual conduct. Nevertheless, as the story goes, he is put to the test by God (and by The Adversary) to see if, after horrible losses and tortures, Job can still keep a good conscience.

As is well-known, what unfolds afterwards is an argument between Job and three of his friends who have come to console him. Job is in a terrible bind: on the one hand, he can not deny that he has faced horrible tribulations, he lost his children, his wealth, his friends and his health. Given that he has never sinned against God and that a fate so horrible could only be an act of God, he has to assert that what has befallen him is unjust. He asserts this through rhetorical questions, even at the risk of offending God, while his friends urge him not to: "By God who has deprived me of justice!" (Job 27:2).

On the other hand, he has to struggle to keep his good conscience and not swear against God. His friends try to ameliorate the situation by claiming that either Job does not have a good conscience and this is why he is punished, or by claiming that God is omniscient and omnipotent and that therefore his actions are inscrutable for us.

What would sinning against God amount to here? In the first instance it would be to curse the name of God. It would also mean to transgress the Commandments. Kant thinks that Judaism conceives of a merely "political" communion with God, and hence sinning in this context might be taken to mean the transgression of the "law of the land": I. e. committing crimes. Job is confident that he has not committed any crimes in the land. Nonetheless, he is being punished, and he can not remain silent in the face of this punishment, since he sincerely believes that he does not deserve it.

It could be objected that "sinning against God" is only possible for the members of a given faith: a faith which strictly prohibits speaking against God. However, the point can be generalized to include non-believers, since it is actually a point about being sincere in ones' religious beliefs. Kant commends Job for nothing other than conscientiousness as he explains the "moral" of the book:

> Hence only sincerity of heart and not distinction of insight; honesty in openly admitting one's doubts; repugnance to pretending conviction where one feels none, especially before God (where this trick is pointless enough) - these are the attributes which, in the person of Job, have decided the preeminence of the honest man over the religious flatterer in the divine verdict. (T 8:267)

As we shall see, these points also resurface in his book on *Religion*, where statutory laws are contrasted with true religious disposition. What matters for Kant is not only the strict

observance of these laws, but rather the disposition with which one follows the laws. The point is important for Kant's overall concern about prioritizing morality over religion. According to Kant we see, in the example of Job, someone who has founded his faith on morality (and not the other way around), whereas Job's friends represent religious flatterers, who seem to only care about the conduct of their lives for gaining what they take to be God's favor.

It is clear that Kant is on Job's side here, and he commends Job for "speaking as he thinks", which is a central feature of conscientiousness. Jobs' friends, on the other hand, speak cowardly, as if they are being judged by a capricious God: for them, gaining God's favor is more important than speaking the truth. What is also commendable about Job is that, towards the end of the book, he realizes that he spoke beyond the boundaries of his understanding and recants some of the things he had said earlier: "Indeed, I spoke without understanding of things beyond me, which I did not know. Being but dust and ashes" (Job 42:3).

What is also striking is that speaking as one thinks and having a good conscience are requirements for the protection of personal integrity: we can see this immediately from the text which Kant quotes: "Till I die I will not remove mine integrity from me" (Job 27:5).

It thus seems that Job is conscientious in three ways: a) he keeps a good conscience-does not sin against God, b) he speaks as he thinks-he asserts that there is injustice in the world even at the risk of offending God and c) he realizes that he spoke about things which he does not fully understand and recants some of his remarks. This

last point is about Job realizing that he has spoken without being subjectively certain, a point that Kant brings up by referring to the famous dictum of Pliny. Generally speaking, then, the *Book of Job* can be considered as one of the first written treatises that allude to something like conscientiousness.

In the text on theodicy we also find an allusion to the role that the analogy of God plays in conscience. Here Kant explains what is expected of conscience in case one swears an oath. Conscience seems to function on this conditional**:** **"**if there is a future judge of the world (hence a God and a future life), the taker of the oath wills to answer to him for the truth of his external profession" (T 8: 269). What is required is not that one profess that there actually is a future judge, but only that one swears the oath as if there were such a judge. In conscience, then, one imagines a "scrutinizer of hearts", as we have seen in the previous chapter.

This passage also suggests that, if I am not sincere in my oath in the first place, being forced to utter some declaration to guarantee my sincerity would not concern me at all. This is why Kant holds that being conscientious in religious matters is more important than following the letter of religious law to the point. In his book on religion, Kant even goes so far to say that if and when humanity is finally educated enough to think freely, true natural religion will become the one and only religion (Rel. 6:68). We also find a similar view in an earlier passage from the same book:

> If a moral religion (to be cast not in dogmas and observances but in the heart's disposition to observe all human duties as divine commands) must be established, eventually all the miracles which history connects with its inception must themselves render faith in miracles in general dispensable. (Rel. 6:48)

The analogy of God, however, could also be misconstrued, if we conceive of God as an agent who is not concerned with our moral disposition, but is rather only focused on the external observance of its dictates. We can thereby be "dispensed from the arduous and uninterrupted effort of affecting the innermost part of our moral disposition" (Rel. 6:169). This is a tendency that is found in people who are excessively strict about external observances, since this can be used to "evade their true moral duty and make up for it by fulfilling the ecclesiastical duty" (Rel. 6:160).

Sticker (2016) and Bacin (2013) bring out the role of the concept of God with regard to the functioning of conscience; according to both of them, the concept of God plays the role of a "morally ideal person". For the reasons stated above, I claim we need to see another use for this analogy, and that is as a "scrutinizer of hearts". The analogy of God represents an agent who knows our deepest commitments and convictions. Not only does it have the perfect conduct, it also has the perfect vantage point from which the reasons for my actions can be inspected from. Here, God is an ideal observer, and we need to form our maxims as if they are being observed by a morally perfect (as well as omniscient) being. This is the manner in which conscience keeps us in awe (respect and fear).

Let us now turn to one of the central issues concerning conscience and religion: the role of God and holiness.

4.2 Divine commands and the highest good

It is no coincidence that we find Kant's definition of religion within his discussion of conscience. Let me briefly refer to a passage from the *Metaphysics of Morals* in which Kant explains this definition and brings out the main connections between morality, religion and conscience:

> [...] the idea (of God) is not given to him objectively, by theoretical reason, but only subjectively, by practical reason, putting itself under obligation to act in keeping with this idea; and through using practical reason, but only in following out the analogy with a lawgiver for all rational beings in the world, human beings are merely pointed in the direction of thinking of conscientiousness (which is also called religion) as accountability to a holy being (morally lawgiving reason) distinct from us yet present in our inmost being, and of submitting to the will of this being, as the rule of justice. The concept of religion is here for us only "a principle of estimating all our duties as divine commands." (MS 6:440-1)

In the first instance, we may ask this question: Why does Kant qualify this sentence with the phrase 'subjectively considered'? One of the most important points that Kant stresses throughout his religious writings is that "recognizing our duties as divine commands" can only be a subjective endeavor, given that knowledge of God or declaring that there is a God can not possibly be required of all human beings (as it can not be an object of experience).

In contrast, the moral principles themselves (which give rise to the duties that ought to be recognized as divine commands) can indeed be shown to be objectively necessary, for all rational beings. Even though it can be shown that moral principles apply equally to all human beings, regarding our duties as divine commands can only be a subjective aspect of our moral lives.

What is the philosophical significance of regarding our duties as divine commands? The point, I shall argue, is about our disposition toward the moral law, rather than our convictions with regard to the existence (or non-existence) of God. That being said, there is a significant role for the concept of God and holiness which we have seen in the discussions of conscience. The point is that we ought to regard the rights of humanity as sacred regardless of our religious convictions.

Regarding our duties as divine commands may help us to register their importance, which is of an incomparably high degree. This disposition toward moral demands could aid us in prioritizing them in our practical endeavors. Kant defines orienting oneself in thinking thus: "when objective principles of reason are insufficient for holding something true, to determine the matter according to a subjective principle" (O 8:136fn). This idea ties in with his account of conscience in the *Metaphysics of Morals*, where the voice of conscience can be considered (subjectively) as the voice of God. By attributing the status of holiness to our moral endeavors, we may obligate ourselves to uphold morality at all costs, together with the appropriate respect that is owed to the demands of morality. This is one of the main reasons why the feeling of awe is taken by Kant to accompany conscience: this feeling reinforces our sensitivity toward what we ought to regard as holy.

In the passage quoted above, conscience is characterized as distinct from us and yet present in our inmost being, as it is taken to be "incorporated into our being" although its commands seem to come from a source other than ourselves. Conscientiousness then turns out to be accountability towards a holy being-which involves regarding our duties as divine commands.

John Hare (2011) argues that Kant fills a "moral gap" between our rational capacities and the demands of morality (a purported gap between ought and can) by relying on or at least hoping for an intelligent author of the world (p. 154). According to Hare, the demands of morality in the Kantian framework are so high that they require "divine assistance" (p. 160). According to another commentator, Patrick Kain (2005):

> Kant's moral theism included the claim that practical reason, reflecting upon the absolute authority of the moral law, should lead finite rational beings like us to believe that there is an omnipotent, omniscient and holy being who commands our obedience to the moral law and proportions happiness to virtue. (p. 139)

I believe Kant does make these claims, but I shall argue that they have nothing to do with "regarding our duties as divine commands". The need for postulating an intelligent author of the world only arises under a certain conception of the "highest good". There is, in fact, a better line of thought in Kant's own understanding of moral religion and conscience. The main feature of this line of thought is Kant's cosmopolitan view of conscience.

John Hare (2011) argues that Kant's moral theism is defensible. In setting up his argument, he relies upon a passage from the third *Critique*, which incidentally contains an allusion to conscience as well. It will be helpful for our purposes to quote it in full:

> Consider a person at those moments in which his mind is disposed to moral sensation. If, surrounded by a beautiful nature, he finds himself in peaceful and cheerful enjoyment of his existence, he feels a need to be thankful to someone for it. Or if, on another occasion, he finds himself in the same state of mind under the press of duties which he can and will satisfy only through voluntary sacrifice, he feels a need to have done something that was commanded and to have obeyed an overlord. Or if he has in some heedless way acted contrary to his duty, although without

> having become answerable to other people, nevertheless a strong self-reproach will speak to him as if it were the voice of a judge to whom he must give account for his action. (KU 5:445)

This passage is found in a section entitled 'ethicotheology'. The strong self-reproach that speaks in the voice of a judge in this passage is an obvious allusion to conscience. Notice that Kant repeatedly talks of a "need" in this passage. I believe this need arises ultimately from the nature of reflexive judgments, in pursuing the notion of a "highest good", which consists in the systematic unity of the ends that we do have (arising from our pursuit of happiness/self-love) with the ends that we ought to have (our duties). Kant foresees this unity as possible only through an omnipotent being (Rel. 6:5).

As I have previously discussed, we use the reflexive aspect of judgment in order to make sense of intentional actions. However, this form of judgment does not warrant us any extension of our theoretical knowledge, and hence it certainly does not warrant us any clue as to the actual existence of God. Kant is very clear that, the fact that we posit an "intelligent author of nature" is contingent upon our "need" to do so. This need, in turn, arises from Kant's idea of the "necessary object of practical reason" which is the "highest good" (KpV 5:135).

The notion of the highest good as the vocation (*Bestimmung*) of the human being can be found in almost all of Kant's critical writings, starting from the *Critique of Pure Reason*, all the way up to the *Metaphysics of Morals* (A801/B829, AP 7:325, P 9:445, KU 5:257, MS 6:437). In the *Critique of Practical Reason*, Kant describes the highest good as "happiness distributed in exact proportion to morality" (KpV 5:110) and as "the necessary object of a will determined by this (moral) law" to which the ideas of God and

immortality serve as conditions (KpV 5:4). The moral law lays before us a possible world in which the highest good would come about if pure reason had sufficient physical power. The moral law commands us to change the world towards the highest good (KpV 5:43).

The highest good is an idea of reason, in Kant's technical use of the term. This idea is "neither intuition nor feeling" but is rather a "concept of a perfection that we can always approach but never completely attain" (AP 7:199). The ideas of reason, then, are not objects of theoretical knowledge, but rather models of perfection. In this sense the highest good can be given as a "plan" which "reason produces according to its own design" (Bxiii).

There are two aspects of the highest good. One aspect is achieving happiness. Kant defines happiness as "the complete well-being and satisfaction with one's condition" (G 4:393) and as the concept of "the sum of satisfaction of all our inclinations" (G 4:399). Furthermore, happiness is pursued by each and every one of us. While happiness can not be the ground of morality (G 4:396), we do have an indirect duty to assure our happiness, since this would make us less resistant to the demands of morality (G 4:399).

The concept of happiness has its roots in our natural inclinations. Kant's account of the affective aspect of human nature has been a topic for debate in secondary literature.[32] While there are passages in which Kant seems to be too harsh in his depiction of inclinations (especially in the *Groundwork*, see G 4:398-9), I believe his

[32] See Formosa (2011), Sorensen (2002) and Hare (2011).

considered view is more moderate and sensible. According to Kant, inclinations are implanted in us by nature (MS 6:277, 6:456; AP 7:152) and by themselves they constitute no evil ® 6:24).[33] Furthermore, we have an indirect duty to cultivate compassionate feelings toward human beings, since the representation of duty by itself might not be enough to incite us to do what morality requires (MS 6:457).

We can find this view also in the *Lectures on Anthropology*. There Kant argues for a middle ground between being consumed by passions and being apathetic. As he puts it:

> Nature implanted in human beings no passions, only inclinations, and it is only fantasy that turns these into passions. Hence, nature also did not want us to observe the apathy of the Stoics with regard to passions and affects. (LA 25:1361)

What is natural, then, can not be evil of itself. Instead, it is free action that prioritizes inclination which can lead to moral evil. This view has obvious ties to Rousseau's thought in general, especially with regards to the broad history of humankind construed as a passage from the state of nature to the state of civilization. Kant grants to Rousseau that the crude state of nature in which inclinations were the only factor determining the behavior of human beings was a state of innocence and happiness in general and that the stage in which we find ourselves is the worst: this is the intermediary stage between nature and civilization which has lead to social inequality. However, according to Kant, the state of nature was also a state of ignorance and hence happiness was only negative,

[33] As were the four aesthetic predispositions of receptivity to concepts of morality, including conscience (MS 6:456). This phrase might sound too teleological to us, however I believe nothing substantial would change if we replaced it with an evolutionary account of feelings.

whereas in a fully enlightened state, we can attain positive happiness and accordingly this ought to be the true vocation of human beings (LA 25:1417-8).

The other aspect of the highest good is the worthiness to be happy. In this respect, the notion of highest good is tied closely with conscientiousness, since we can only be worthy of happiness if we act sincerely according to our best judgments, which in turn have to be observant and obedient towards the moral law.

There seem to be two general problems involving the highest good. If we construe the highest good as involving the actual attainment of happiness, we set the standard too high (Rel. 6:182, KpV 5:129). There is a fundamental vagueness about the concept a highest good in itself, given that Kant construes happiness as indeterminate (G 4:418). There doesn't seem to be a reliable way of translating moral worth into reward in the form of happiness, since the conception of happiness for each and every person is bound to be different/incommensurate. The second problem is more fundamental, and it relates to Kant's Transcendental Idealism. Given the gap between noumena and phenomena, there arises an irreconcilable divide between nature and free action. This is why only an omnipotent being can guarantee the achievement of the highest good.

In this connection, I believe there are two ways of making sense of the highest good. The first is "the existence of rational beings under moral laws" (KU 5:444) which is only to construe the highest good as a normative goal by leaving out the actual attainment of happiness. The second way is to incorporate the actual attainment of the highest good, and this is the view that I find problematic for the reasons stated above. John Hare and Patrick Kain construe the highest good from this perspective, which is

why they emphasize Kant's moral theism. Their argument also rests on a strict understanding of the principle of "ought implies can". If our powers are insufficient to bring about the attainment of happiness through moral means, how, then, can we be obligated to do so?

The 'ought' implied in the highest good, however, must belong to a different kind of order than our specific moral duties. In the first instance, the highest good is an ideal and it that sense it is by definition impossible to achieve completely. In this sense, it is entirely different from our duty to refrain from deceiving others. Regarding the highest good, we can only be obligated to strive for the worthiness to be happy. The attainment of happiness, by definition, is not a duty, since we already strive for it due to our form of agency.

The highest good, then, can be helpful for giving an orientation, in the sense of the promotion of our endeavors with a view to worthiness. However, this concept is not necessary for the determination of our regular duties, instead, it grows out of our reflection upon the moral law and its extrapolation into the natural order of things. Thus, if we were to show that the highest good is achievable only through human means, or if we construed the highest good as requiring only that we do all we can in order to be worthy of happiness, the postulate of God would not be necessary.

What is especially significant in this connection is that, throughout his writings, Kant consistently holds the view that morality has priority over religion. The chief role of religion is to make the concepts of morality more amenable to sensibility and hence it has to be regarded as part of a moral anthropology. The anthropology is crucial, but it

has nothing to do with the justification of moral, hence it does not (and can not) play the role of a measuring rod, but only the role of a mainspring. As Kant puts it in the *Groundwork*, "[e]ven the Holy One of the Gospel must first be compared with our ideal of moral perfection before he is cognized as such" (G 4:408).

Hence, anyone who recognizes the demands of morality is under the obligation to uphold them, regardless of faith or conviction. As we have seen, Kant thinks that any person who possesses moral principles is endowed with a conscience (VL 27:574). The sole requirement of morality is that the person be rational; which entails the capacity for recognizing the demands of morality. Our conscience, in turn, can be used to ascertain whether we are in fact worthy of happiness, since it is concerned with our sincerity.

Patrick Kain (2005) discusses a passage from the *Doctrine of Virtue* in order to argue that, in Kant's view, "the existence of God is inseparably linked to the 'binding' or 'obligating force' of morality" (p.130). I believe this idea is fundamentally mistaken. In that passage Kant states that "we cannot very well make obligation (moral constraint) intuitive for ourselves without thereby thinking of another's will, namely God's (of which reason in giving universal laws is only the spokesman)" (MS 6:487). However, just in the previous sentence Kant had remarked that "the ground on which a human being is to think of all his duties in keeping with this formal aspect of religion (their relation to a divine will given a priori) is only subjectively logical".

What is of utmost importance is that the binding force of morality precedes our recognition of duties as divine commands. We can see Kant making this point as early as in the first *Critique*:

> So far as practical reason has the right to lead us, we will not hold actions to be obligatory because they are God's commands, but will rather regard them as divine commands because we are internally obligated to them. (A819/B847)

The importance of regarding our duties as divine commands is not related to the obligating or binding force of morality, but is rather related to our disposition towards the demands of morality. This disposition is about upholding the demands of morality above all else, giving it our utmost respect (since it is inviolable and incomparable), and determining the worth of any practical endeavor only by recourse to the moral law.

This point is evident throughout Kant's critical writings. In the *Religion*, even the observance of legitimate civil laws is considered to be a divine command (Rel. 6:100fn). What Kant regards as sacred is the rights of humankind. Morality itself is sacred, prior to any introduction of religious doctrine. Thomas Green (1998) makes a brilliant point about why some things ought to be regarded as sacred: the opposite of sacred is the banal (p. 112).

We do need to regard the demands of morality as sacred, inviolable, incomparable, but this does not commit us to any religious doctrine. It is not because we are rationally committed to the belief in God, but rather the duty to be truthful in a case having to do with what is most sacred of all among human beings (the right of human beings) (MS 6:304). Moral commitments entail an affective engagement, no matter our religious convictions. Part of this affectivity is reflected in our conscience, which keeps us in check through feelings the feeling of awe (MS 6:438).

This point is related to the cosmopolitan view of conscience. According to Kant, natural religion is "that religion in which I must first know that something is duty before I can acknowledge it as a divine command" (Rel. 6:154). The religion of reason is cosmopolitan, in that it "will have all right-thinking human beings as its servants" (Rel. 6:152-3). To be truthful is also a "sacred command of reason (SRL 8:427). Even the idea of a civil constitution is declared to be sacred by Kant (MS 6:372). The divinity of commands, then, is consequent upon their legitimacy. It is evident from these passages where Kant stands on the Euthyphro question.[34]

Contrary to Patrick Kain (2005), I do not think that there is a direct connection between regarding our duties as divine commands and the highest good (p.139). The main reason is that Kant's definition of religion does not refer to the possibility of happiness, but rather emphasizes the correct moral disposition that we ought to have with regards to our duties. Kant's postulation of God's existence relies solely on his conception of the highest good, insofar as we incorporate the actual attainment of happiness into that conception, but this has nothing to do with the holiness or the binding force of morality. To claim that they do would violate Kant's own cosmopolitan approach to religion and conscience.

[34] For an extended discussion of the Euthyphro question with regards to Kant's critical philosophy, see Westphal (2016).

4.3 Conscience in the *Religion* (1793)

In a letter written to Staudlin, Kant introduces his book on Religion in this way:

> Conscientiousness and true respect for the Christian religion have been my guide in this work, but also the principle of a befitting candor: to conceal nothing but rather to lay open how I believe that I see a possible union of the Christian religion with the purest practical reason. (May 4,1793; AK 11:414)

At the outset, a series of objections comes to mind: Why do we require such a union at all? Wouldn't another religion also suffice for the "subjective aspect" of morality? What makes Christianity so suitable for a possible union with pure practical reason? Reading onwards in the text, we realize that Kant is well aware of these objections, and he opts for a distinction which might today sound peculiar: a distinction between faith and religion.

According to Kant, whereas there are many different faiths (Christian, Mohammedan, Judaic, Indian and so on) there can only be one religion (the recognition of our duties as divine commands). The idea here is that, given that the necessary transformations within their respective doctrines are carried out, any faith can come to be in a union with pure practical reason. The most basic principle required is that a given faith be accommodated such as to be based on morality.

According to Kant, the rational development and interpretation of these faiths will gradually lead them to dispense with requirements for belief in miracles, since the authority of duty will suffice for the strength of our moral disposition. Kant's idea is that these miracles were needed to illustrate the holiness of moral commands and convince

people of their validity and authority (Rel. 6:84). They were merely moral aids, designed to improve the moral disposition of believers. Through natural religion, there will be no more need for such aids, as we shall recognize our duties as divine commands.

What is important is that this change is a change in moral disposition, but this also counts as a deed (Rel. 6:74), in that the person is fully responsible for the change, and no other kind of atonement (e.g. sacrificial offerings, ceremonies) can take its place. Kant refers to a passage from the New Testament to explain the birth of a new man through a change of heart (Rel. 6:47):

> Do not lie to one another, seeing that you have stripped off the old self with its practices and have clothed yourselves with the new self, which is being renewed in knowledge according to the image of its creator. In that renewal there is no longer Greek and Jew, circumcised and uncircumcised, barbarian, Scythian, slave and free; but Christ is all and in all!". (Colossians 3:10-11)

The passage is striking for its cosmopolitan tone. This cosmopolitanism could be one of the reasons that Kant claims Christianity is the religion most suited for a union with practical reason. What Kant has in mind is a "purification" of religion from what he calls historical and statutory practices (or cultic observances). What is left after this purification would be, for Kant, true natural religion which has a very basic definition: "Religion is (subjectively considered) the recognition of all our duties as divine commands" (Rel. 6:153-4).[35] Here Kant gives us the hint to his overall strategy of biblical exegesis: to move from revealed religion into a natural/moral religion. The difference between them is the priority given to divine commands and to duty. In revealed religion, divine commands are necessary in order to recognize duties. In natural

[35] I shall discuss this definition of religion in the last section of this chapter

religion, by contrast, we know that something is a duty before recognizing it as a divine command. So, in natural religion, morality precedes and grounds the dictates of theology and, concurrently, our interpretation of religious texts should aim to give the same priority to morality.

Bringing together the subjective aspect of morality (the recognition of duties as divine commands) with the objective aspect (moral law as given by pure practical reason) constitutes the union between religion and practical reason. Throughout the *Religion*, Kant construes Christianity as the most suitable religion for such a union, and the concept of conscience has an indispensable role to play here.

As Onora O'Neill points out, the exposition in the *Religion* follows "a traditional Christian articulation of human origins and destiny: original sin, temptation, conversion, and ministry are moments of the encounter of the pilgrim soul with good and evil" (O'Neill 1996, 291). The reason for this could be that Kant's audience is "eighteenth-century Lutherans, who already accept the Christian doctrine of original sin." (Wood 2014, 56).

Throughout the book, Judaism is contrasted with Christianity. The former is found to stand "in absolutely no essential connection i.e. in no unity of concepts", based on a theocracy with "no rights over, or claims upon, conscience" (Rel. 6:125). By contrast, Christianity can propagate more easily among different peoples and in different times since it carries not a statutory but a moral message, thereby treading in "close proximity" to reason (Rel. 6:167).

According to Kant, this is one of the most important distinctions between Judaism and Christianity. In the former, God is represented as a secular regent tied only politically with its believers, whereas a universal moral motivation comes to the fore in the Christian religion. This is not to say that Kant finds the entirety of Judaism useless. On the contrary, the Ten Commandments, for instance, have ethical validity. What is missing, from Kant's point of view, is an account of (and claims upon) a moral disposition with which these commandments are to be followed (Rel. 6:126). As mentioned in the previous section, this is quite peculiar, especially in lieu of Kant's discussion in the *Theodicy* essay about the story of Job.

Now, the possibility of a transition from revealed to natural religion is already inherent due to our physical and moral predisposition (*Anlage*). This transition amounts to religion being "gradually freed of all empirical grounds of determination, of all statutes that rest on history and unite human beings provisionally for the promotion of good through the intermediary of an ecclesiastical faith" (Rel. 6:121). Kant argues that this kind of progression already took place, when Judaism, influenced by Greek wisdom, eventually gave rise to Christianity (Rel. 6:128).

Commands that have a historical source will always remain dubious and contingent, especially when compared to the commands of reason. What is required, therefore, is the maturation of humankind, by which we recognize our duties (issued by reason) as divine commands (this is the definition of religion).

Kant's worry here is that any kind of person can follow the letter of the law, and if religion were merely about performing certain rituals and observances, anyone could

be religious. What is of utmost importance, however, especially with regards to religion, is that a person is honest with him/herself. In this sense, sincerity is "the foundation of conscience and all inner religion" (Rel. 6:190).

A further idea here is that observances themselves could be irrational and hence biblical exegesis must be firmly based on morality. In this respect, Kant's interpretation of the story of Abraham sharply contrasts with his interpretation of the *Book of Job*. Here, Kant is strictly opposed to Abraham's conviction and conduct. Indeed, in the *Conflict of the Faculties* we see how Kant interprets the whole situation: "Abraham should have replied to this supposedly divine voice:'That I ought not to kill my good son is quite certain. But that you, this apparition, are God - of that I am not certain, and never can be, not even if this voice rings down to me from (visible) heaven' (CF 7:63fn). According to Kant, there can not be a divine imperative which orders an immoral deed.

He makes a similar claim in the *Religion*. Even when we are faced with what is seemingly a divine miracle (God speaking from heavens), we need to make sure that what is commanded is in keeping with the moral law. The source of an immoral command "cannot be a divine miracle despite every appearance of being one" (Rel. 6:87).

Recall that our conscience functions as a fictional person (or deity) who has access to our motivations and maxims. It is indeed pointless to feign conviction in the face of someone who can "see through" us. We imagine the voice of conscience to belong to such a person, this is what makes it effective. That is, we imagine a person or a being who has access to our inner thoughts; this is someone we can not possibly deceive.

Hence, in true Kantian fashion, we ought to take special care to act only on those motives and maxims which can be made public/communicable-hence they must be communicable and also acceptable (in principle) by our peers.

As Onora O'Neill (1996) explains, what this requires is that our thoughts and actions are organized as "lawlike", since otherwise they "will be unfollowable by at least some others, who will view them as arbitrary or incomprehensible" (p. 276). This point has obvious allusions to Kant's essay on Enlightenment. Our motives, maxims, reasons, ends and actions have to be available for public scrutiny, and we must engage critically with the moral/political thoughts and actions of our peers.[36] This is one of the many places in which morality and politics might come together. In that sense, the concept of conscience is especially important in bringing together virtue and justice. With regards to religion, then, we must follow the same logic. The way to this is by putting forward a moral religion which does not require any dogma but rather requires that our actual duties are regarded as divine commands.

In Kant's view, this is the way in which previous religious commentators or rational theologians have always treated the articles of faith, as did the ancient people with their mythology (Rel. 6:110-1). The idea is that the interpretation of religion and mythology were ultimately aimed at the moral cultivation of human beings, as they ought to be. It is safe to assume that Kant sees himself as another one of these "rational and thoughtful" commentators and that his book on *Religion* is a continuation of this tradition: a moral hermeneutics. What we find in this book is an evolutionary account of

[36] See (Rel. 6:134) in which Kant ties together the importance of free thought with the necessity of free speech.

religion, starting with the ancients and their mythology, going through Judaism and Christianity and ending in natural religion (which is brought about by enlightenment) (Rel. 6:111, 6:121fn).

Religion, then, can play a role in the moral cultivation of agents. One way in which this is possible is through an analogy of a lawgiver for all human beings (MS 6:440). Why do we require the "analogy" at all? Why is the law itself not sufficient? One important point is that the analogy is not about the content of the law, but with our relationship with law. Mere compliance with the moral law can not be the ground for moral worth, we need to do the right thing because it is the right thing to do. This requires the right kind of relationship with the moral law, and establishing this relationship is not possible without the concept of conscience, since it is concerned both with imputability and with our receptivity to concepts of duty via the feeling of awe.

In this context, I believe the *Religion* text could be construed as part of Kant's moral anthropology, as it incorporates the idea of subjective ends with the idea of duty. In other words, the *Religion* text is another attempt to build a connection between rational law and human existence. Kant sets out our final end as "good life conduct" which is also the goal of becoming "pleasing to God" (Rel. 6:85).

Here, Kant is not claiming that we have a certain connection to God and this is why we ought to act in such a way as to please him. The phrase is rather about putting forward a standard for which to strive for in our moral endeavors. This is precisely the way in which we ought to interpret the voice of conscience as the voice of God: we must hold ourselves to the standard of a morally perfect being. This standard can not merely

be compliance with the law, but it must be our observance of it with due respect. Furthermore, one must "constantly test himself as if summoned to accounts before a judge" (Rel. 6:145).

The ideal of a perfect being is always bound to rouse a tension within us, as we compare our conduct with what we consider to be a morally perfect conduct. In cases where we are able to approach this ideal, we shall feel relief and in cases where we stray away, we are faced with pangs of conscience. This is in line with Martin Luther King's idea of a tension that is necessary for moral improvement.

This tension arises because, as rational beings that belong to the world, our own happiness is inevitably our most basic and final end (Rel. 6:7). In addition to this, we ought to make the highest possible good in this world our ultimate end as this is commanded by pure reason.[37] Obtaining the highest good in this world is about promoting the happiness of our fellow human beings, apportioned to their worthiness to be happy. The "wide duty" that we have of making the happiness of others our own end only extends to those who are worthy of happiness; and to be worthy of happiness is precisely about doing the right thing because it is the right thing to do. Conscientiousness, since it urges us to act according to our best judgment and with the best intentions, is crucial in ascertaining our own moral worth.

[37] The end that contains the inescapable, and at the same time sufficient, condition of all other ends is the final end. One's own happiness is the subjective final end of rational beings belonging to the world (they each have this end by virtue of their nature which is dependent upon sensible objects; it would therefore be otiose to say of that end that one ought to have it), and all practical propositions that have this ultimate end as their ground are synthetic yet at the same time empirical. (Rel. 6:6fn)

Alas, our capabilities and resources are limited. We can not possibly "effect happiness in the world proportionate to the worthiness to be happy". For this reason, we can only hope that "an omnipotent moral being" would bring this about. It is through this reasoning that Kant claims that **"**morality leads inevitably to religion" (Rel. 6:8).

At the outset, it is not clear why this transition is "inevitable". Is it not enough for us to do our best in the pursuit of this final end? Is it not meaningful in itself to struggle for the highest good, even though we know our contribution will be infinitesimal? A parable found in some Talmudic and Islamic resources portrays such an understanding of morality. In this parable, Nimrod lights a huge fire in Damascus and wants to cast Abraham into it. A crow is carrying bushes to the fire, while an ant is carrying a drop of water. The crow, mocking the ant, asks what good would a drop of water do against this gigantic fire. The ants' reply is this: "at least I'll show where my allegiance lies". The ant is aware that her effort will not result in a significant change in the situation, nevertheless she does not give up the struggle.

The parable seems to question one of the main principles of morality in general. Kant's main concern in this context can be thought of as related to this principle, the principle that "ought implies can". According to him, if we are to work towards an extremely difficult task, we need to have at least some hope that it is possible to attain our goal. Here, Kant is not thinking about an incentive or a promise to turn us towards morality; rather, he is seeking a ground that warrants belief in the successful completion of a seemingly impossible moral endeavor, which is given to us clearly by reason. Kant holds that promoting the highest good can not be a duty for us if we believe it to be impossible.

O'Neill (1996) stresses the importance of the dictum "ought implies can" in this way: "[t]hose who think action that changes the future impossible can aim for nothing: commitment to action that is thought impossible is not really commitment; we cannot aim to achieve what we know to be unachievable" (p. 284). However, some acts of civil disobedience, of which a paramount example of was witnessed in Tiananmen square in 1989 (the anonymous man facing the tanks), seem to exhibit instances where people commit themselves to challenging some power or authority upon which they can exert minimal change (or none at all).

Obviously, these are extreme cases, and we can not build our moral system upon these cases. This point is not about morality per se, but about rational hope and volition in general. It simply makes no sense to hope for an impossible outcome. On the other hand, where we stand, morally and politically, means everything to us, even if we are in what appears to be a "losing battle". The ant in the parable is not hoping that the water drop it carries will extinguish the fire. She might well know that this is impossible.

What is at stake here, then, is personal integrity, as it was in the discussion of the *Book of Job* in the *Theodicy* text. Job was also struggling to keep his faith in the face of insurmountable agony. He was desperate, he had no power to change the course of his life, and yet he maintained his integrity. He kept to his word, regardless of the circumstances in which he was in. Kant even commends Job for keeping his integrity. Hence, we could argue that there may be cases in which our sense of justice (a central notion of our personal integrity) might come into conflict even with the principle that "ought implies can".

The relation between personal integrity and conscience can be conveyed with Luther's famous objection[38] against the Diet of Worms: "Here I stand, I can do no other". This exclamation both expresses a demand from authority to provide a space in which one is able to think and live in accordance with one's own conscience; and also, it is a demand that we impose upon ourselves in order to sustain our personal integrity by acting/thinking in the way we sincerely hold to be morally (or religiously) accurate. Going against conscience results in the rejection of "both the authority of the moral law and [ones] moral identity as a person" (Bacin, 2013, p. 253). As I have shown, Kant argues that sincerity is the basis of conscience and a lack thereof would turn everyone into "inward hypocrites" (in this connection, see Rel. 6:190fn).

Kant also draws attention to the importance of personal integrity during his lectures. As he puts it:

> Uprightness, and that in its greatest purity, i.e., integrity, candor et sinceritas, are natural obligations of man, and so everyone must frame only such utterances as can coexist and agree with the greatest consciousness of truth and the total absence of any consciousness of the opposite (VL 27:699)

While integrity is certainly significant, it can not be our only criterion. After all, we can think of evil people who possess personal integrity. The conception of integrity discussed here points to the agent being adamant in following what she holds to be the right thing to do, even in the face of extremely difficult, oppressive or straight-out violent circumstances. In other words, it concerns her best judgment, and hence her

[38] Whether he actually uttered this sentence in these exact words is debated, but this does not matter for the point I want to make.

conscience. It is always possible that we might be mistaken in what we hold to be the right thing to do. We might even be irrational, in the sense of being imprudent or naive. We might insist on fighting a losing battle, no matter the cost. In some instances this might make us heroes of conscience. In other instances, all we might get could be misery and desperation.

What we need on top of integrity, then, are the right moral principles and the right approach to judgment, on top of the cultivation of conscience. In other words, we need the correct measuring rod and this requires something more than subjective certainty. Arriving at the right moral principles and acting on them is possible only by a comprehensive moral education, in which the pupil learns not only the rules of morality but also the practice of judgment. I shall return to this issue in the next chapter.

One central idea of the *Religion* is the maturation of individual persons and humanity as a whole. This maturity (which Kant also calls enlightenment) requires that we become not merely legally good but also morally good (pleasing to God). The point can be clarified by using the terminology that Thomas Green employs: mere compliance with norms is not yet conscientiousness; what is required is the "observance" of norms (Green p.33). That is, accepting norms as genuine standards for our conduct and employing them to the best of our abilities.

Again, why is mere compliance with the moral law not enough? One general answer could be construed in this way: The more we tend toward maturation, the less we regard the moral law as an external imposition upon our will, i.e. an imperative. We come to see the necessity of the moral law not as an something forced upon us, but

rather as the true ground of free and responsible conduct. It becomes a part of our moral constitution to do the right thing because it is the right thing to do. This disposition gradually reinforces our personal integrity, since we tend to follow the dictates of our own reason.. Hence, the importance of becoming not only legally good but also morally good concerns our sense of personal integrity together in conjunction with the utilization of correct moral principles.

Kant claims that at this point, after we have been educated to maturity, we no longer need religious rituals and observances, since, as a species, we have become adequately moral by internalizing the requirements of morality as central to our moral character (Rel. 6:123-4). Maturity, then, is not merely about following the letter of the law, but about how we are disposed in the face of the law. A progressive story could be told here: the more enlightened we are, the less the moral law seems like a burden or imperative. In Kantian terminology, the more we approach a holy will, the less the moral law seems to carry an imperative. When the moral law ceases to be a burden, the human being will incorporate his/her maxims according to his/her best judgment much more easily. Hence, maturity is tantamount to conscientiousness.

Kant concludes the *Religion* with a discussion of conscience. The points and arguments he puts forward here are in line with the Vigilantius lecture notes and the text on Theodicy. It is also consistent with what he wrote in the *Metaphysics of Morals*. However, whereas in the *Metaphysics of Morals* conscience was considered from a broadly moral perspective, in this text there is much more emphasis on the religious aspect of conscience, as can be expected.

Once again, in this last section of the *Religion*, Kant is careful in distinguishing conscience from the determining use of the power of judgment. Judging whether an action is right or wrong is not in the purview of conscience (Rel. 6:186). Conscience is concerned, instead, with the degree to which we are certain as to the moral status of a proposed action. What is at issue here is subjective certainty, as it was in the Vigilantius lecture notes.

In conscience, Kant writes, "reason judges itself, whether it has actually undertaken, with all diligence, that examination of actions, and it calls upon the human being himself to witness for or against himself whether this has taken place or not" (Rel. 6:186). Characterizing conscience as a witness here seems to foreshadow the court metaphor in the *Metaphysics of Morals*, where conscience was deemed a "consciousness of an internal court" (MS 6:438). Again, the emphasis is on the reflexive aspect of judgment where we ought to examine carefully, with due diligence, whether we have in fact made the best judgment possible, given the facts about a particular moral situation.

In this part of the text we can see that Kant's arguments for the infallibility of conscience carries a political import. By arguing that there can be no sufficient grounds for certainty with regard to religious convictions (since all our knowledge of revealed religion come from historical sources and hence can be in error), Kant arrives at the conclusion that all religious persecution is strictly impermissible (Rel. 6:187.) Accordingly, a person who has a different religious conviction (or none at all) cannot be blamed for lack of conscience: instead, those who persecute people for having unorthodox religious belief are themselves acting unconscientiously (Rel. 6:187-8).

What makes religion a moral issue is that any assent to a religious doctrine has to be made freely. An authority which attempts to impose a certain religious view is, then, by definition immoral, since it attempts to keep its subjects in a state of immaturity. This kind of oppressive religious authority, however, is not the only target of Kant's criticism: becoming mature agents requires thinking for oneself and reaching ones own conclusions, especially with regard to faith. Kant had stated his view of this issue as early as 1784 in his essay on Enlightenment, where he wrote that having a "spiritual advisor as a conscience for me" is tantamount to remaining in a state of "self-incurred immaturity" (WE 8:35).

Kant reinforces his view in the last section of the *Religion* by pointing towards the importance of the cultivation of a more liberal attitude toward religious convictions. He argues against some religious authorities who claim that some group of people are "not yet ripe for freedom" by stating that thinking in this way would force us to remain unfree, since "we do not ripen to freedom otherwise than through our own attempts (and we must be free to be allowed to make them)" (Rel. 6:188fn).

Kant takes on another form of argument which is designed to force people into believing in a certain doctrine of religion: this is the famous wager put forward by Blaise Pascal in his *Penseés*. The wager can be stated in this way:

> a) Either there is a God, or there isn't.
>
> b) If there is a God, and you believe in Him, you shall be rewarded with a blissful eternity in heaven.
>
> c) If there is a God, and you do not believe in Him, you shall be punished for eternity in Hell.
>
> d) If there is no God, and you still believe that there is one, this will only cause you minor inconvenience.

> e) The most rational course of action is to believe in the existence of God, simply in view of a risk-calculation. (p. 154)

Kant explains precisely why this wager fails to grasp the significance of religion and conscience. If I profess to a belief that I do not hold sincerely, merely because of an anticipated risk, I am in effect believing dishonestly. This results in making religion merely a prudential issue, rather than a moral one. Kant proposes instead that we ought to act in a manner that is pleasing to God, which means acting in accordance with the principles of morality, whether we in fact believe there to be a God or not: the safety before conscience can only consist in this disposition (Rel. 6:189).

In this connection, conscience is once again tied closely with sincerity, which is the attitude that "everything said be said with truthfulness" (Rel. 6:190fn). The cultivation of this attitude is of primary importance. I shall return to this issue of cultivation of conscience in the next chapter.

The final issue to discuss here is the role religion plays in representing what is invisible (regarding all true duties as divine commands) in something visible. Kant presents four religious practices which can serve the good if accompanied by the appropriate moral disposition and sincerity: a) private prayer-which "awakens our heart to the disposition to the good", b) public assembly as a way of propagating the correct principles and thereby sharing them, c) transmitting the principles into posterity by admitting new members into the congregation and d) the maintenance of the unity of this congregation through repeated formalities (Rel. 6:193).

Regarding Kant's whole argument in the *Religion*, of the unification of practical reason with the tenets of Christian religion, I believe these practices ought to be regarded as having a universal scope under a cosmopolitan principle. The main reason for this is Kant's insistence on the freedom of conscience considered as sincerity in religious belief. If this community were restricted only to the members of a specific faith, Kant's idea of a moral religion would be lacking in universality and conscientiousness. Furthermore, Kant foresees a lesser need for these practices in case we truly improve our moral dispositions, that is, the more mature and rational we become, the less we require the visible representations of a communion. These visible representations play a significant role in moral education and it is to this aspect of Kant's thought that I now turn.

CHAPTER 5

MORAL EDUCATION, CONSCIENCE AND CRITIQUE

So far I have discussed the main aspects of Kant's philosophy, his conception of conscience and its relation to religion. I shall now focus on issues regarding the distinctive quality of morally salient features in our lives and see what role conscience might play in the formation of character through moral education.

The reason we require a specifically moral education is that moral facts are qualitatively distinct from any other kinds of facts. This feature of moral facts can be learned not by reading books or by memorizing propositions, but rather through practice and deliberation. As Barbara Herman (2005) explains:

> We do not see moral facts in the direct way that we perceive colors and shapes. We do not grasp the moral truths about things by being informed of their names and natures. We require certain experiences—moral experiences—and interpretations of the experiences (instruction) to become aware of and responsive to a moral world. (p. 257)

The rightness or wrongness of a particular moral action strikes us as so due to our recognition of these moral features. As I have discussed in the second chapter, the ability to recognize actions as morally significant presupposes the reflexive use of judgment through which we attribute intentionality to agents. In this connection we could say that there is a sense in which moral facts are intelligible and not merely empirical (recall the chess example from the second chapter). As we have seen, Kant also stresses the importance of comparing our actions with our moral ideals through the use of

conscience. By engaging in moral situations and acquainting ourselves with various and diverse examples, we come to form moral principles, commitments, preferences and practices which all contribute to the formation of our character.

The formation of character concerns acting from principles and with resoluteness (P 9:487). Resoluteness requires being committed to a certain principle and following it as far as one can. Since the concept of conscience is tied to imputability (discussed in the third chapter) and personal integrity (discussed in the fourth chapter), I believe it should also have ties to our commitments to moral principles. Commitments, then, are central to what we take to be valuable, dignified, right and even sacred. A person committed to the "sanctity of human life" for instance will tend to prioritize this feature in moral deliberations in which this commitment becomes relevant. Mature and educated judgment requires that we demarcate which principles are relevant to what kinds of cases, as we tend to view moral issues through the lens of the moral commitments we happen to have.

Commitments do not only pertain to morality, but can also be held with regards to religion, craft, political or economical views and so on. A person can be committed to providing the best kind of result that she can in her profession, and another can be committed to taking the easy way in practical endeavors. There will also be degrees of commitment, some of them will be more amenable to modification while other commitments will be more central to our self-conception and hence more difficult to change. These latter kinds of commitment can be called our core commitments and they are constitutive of our personality.

The problem is that our core commitments are rarely open to revision and assessment, since they are partly constitutive of our moral character and hence are elementary for personal integrity. Furthermore, we can not function as moral agents through constant struggles to determine what ought to be our most fundamental commitments, as this could lead to a sense of disintegration or suspension of character and hence might be morally harmful.

For this reason, we require a degree of consistency and stability of character, both in ourselves and in others, so that our social interactions are intelligible and that we can form reliable expectations with regards to each other accordingly. Thinking and acting intelligibly, then, requires a certain alignment between what we expect would happen and what actually does happen. Acting intentionally presupposes some awareness of the consequences of one's purported actions. We hold people responsible for their intentional actions partly because we expect them to also be aware of the consequences of their actions.

Commitments, then, are central for any morally significant interaction. However, there is the danger of having the wrong kind of commitments (e.g. being committed to the advancement of one's own interests to the detriment of the conduct of other people). This is why having the correct moral principles to evaluate and assess our core commitments is just as important as acting in accordance with our commitments. The moral law, then, has to provide objective constraints upon any and all forms of action, and also upon our commitments and hence character as a whole.

In this context, conscience can be construed as our sensitivity toward our commitments, especially in moral, religious or political situations. Typically, we tend to experience a violation of our core commitments as a direct assault on our personhood. This violation could be something that happens around us, or it might be our own doing. As we have seen, Kant also construes the violation of conscience as tantamount to the loss of one's moral worth (VL 27:575).

Some commitments, however could be so morally perverse that we could have a duty to confine those who embrace them. The criteria which decides whether a commitment is morally permissible should not refer to any commitment that an agent might happen to have: it should instead refer to a commitment that all agents ought to have. The measuring-rod called the moral law, then, can and ought to be used for moral criticism, not only in the assessment of actions but also in the assessment of character.

Thomas Green (1999) explains two distinct types of voices, involved in conscience, that relate to internal and external criticism. External criticism is about getting the correct moral principles to bear on our choices, actions, policies and maxims. Internal criticism is about "being the person that we claim to be", that is, holding ourselves to the standards that we have set for ourselves, whether in terms of honesty, generosity, integrity and so on. Internal criticism pertains to integrity and how we come to view or interpret it, as it speaks of "truth in the logic of coherence", therefore its claims of falsehood are claims of "dissembling" (p. 60). Dissembling in this context is a sense of personal disintegration.

Mature and rational agents are those agents who are sensitive to both kinds of criticism: they want to do the right thing for the right reasons, and they want to act consistently. Mark Timmons (2002) points to two distinctions with regard to right action: the formal vs. the material rightness of an action and the objective vs. the subjective rightness. Formal rightness concerns the first person understanding of a situation: "the act in question is the right to perform in situations that are such as the agent believes them to be" (p. 259).

Material rightness comes about if the agent's beliefs about the situation are in fact correct. Objective rightness amounts to material rightness. However, an action is subjectively right if the agent believes it to be objectively right (ibid. p. 260). These distinctions are germane to the conception of conscience that I have discussed. The best judgment that conscience informs us of are judgments which we have come to believe to be subjectively right, and hence we take them to be formally right. It is in this sense that conscience can not be in error. Material or objective rightness, on the other hand, can not be solely determined from the first-person perspective. This is why acting in accordance with our conscience can still lead us to moral error.

Moral education involves a tension between these two standpoints of internal and external criticism. On the one hand, the pupil must learn to form his/her own judgments in moral situations (as well as his/her moral commitments in general) and has to do so with a sufficient degree of confidence in order to be functional in society. On the other hand, s/he must always keep in mind the possibility of being mistaken and hence must be sufficiently self-critical. Emphasizing any of these aspects too much can result in

moral stultification, the pupil may become too self-confident or too-insecure. Mature agents are those agents who can form a balance between these two standpoints.

What is required of a moral education (or any education for that matter), then, is that the pupil learns both the limitations and the ranges of his/her cognitive powers. A sound grasp of these features will help the student to see what s/he is capable of in the assessment or criticism of a particular situation, and where his/her capacities become insufficient. In this way the pupil can come to act and think with confidence while remaining open to suggestions and criticisms.

Here are some of the central questions that relate to moral education: what are the characteristics of that being which is to be educated? What are its abilities and limitations? What methods will prove to be effective in educating that specific being? Finally, how does conscience fit into the project of moral education? Answering these questions requires looking at Kant's anthropology, pedagogy, as well as the sections entitled "doctrine of method" in all three Critiques, and the *Metaphysics of Morals*.

5.1 Conscience and critique

Let us begin with the link between conscience, critique and moral education. There is a sense in which a "critique" of reason is an inherently conscientious endeavor. Kant argues that we have a "duty of self-examination" (MS 6:438). This self-examination is to be carried out without recourse to any external authority other than our own reason and hence is reflexive. This is apparent in Kant's repeated reminder that we ought to

compare ourselves not with other people but with the moral law itself (as I have shown in chapter three).

Now, what is involved in a "critique" of any faculty, in the Kantian sense? As Allen Wood (2007) explains, "a faculty is the way that a living being achieves something through processes or actions that are normatively conceived and normatively guided" (p. 115). If we look at Kant's critical writings we shall see that they are often examinations of these processes and actions which determine the principles that guide our rational capacities (or at least principles which "ought" to guide them). These writings, then, can be used to teach the pupil of moral education about the limitations and ranges of his/her cognitive powers.

Since all faculties are normatively guided, all practice that we engage through these faculties also involve norms. Practice is not the same as mere activity, behavior is not conduct and a response is not a mere reaction. For this reason, Kant defines practice normatively: "not every doing is called practice, but only that effecting of an end which is thought as the observance of certain principles of procedure represented in their generality" (TP 8:275). As Thomas Green (1999) puts it, the aim of education is the "the effective regulation of conduct" via "self-governance" (p. 3). Indeed, this is a Kantian point, as the cultivation of character is about "acquiring an aptitude of acting according to maxims" (P 9:481).

Self-governance will involve our obedience to certain norms that we take to be authoritative. We can only employ these norms reflexively, that is, by examining whether our conduct lives up to the standards that we set for ourselves. These personal

standards and commitments, fallible though they may be, are reflected in the responses of our conscience. As we have seen, conscience is about self-examination which aims to discern whether we have acted (or intended to act) according to our best judgment. The Kantian lesson, then, is that ultimately the moral law must be used as a final arbiter, the authoritative measuring rod for the rightness of our actions and commitments. If our commitments are in conflict with the moral law, they have to be amended.

In correlation with this point, our best judgments have to be those judgments which we believe to be justifiable. Deciding what we can or can not justifiably assert in any domain (whether morality, politics, religion or even science) is not only a task of epistemology, but also a social and moral requirement. The only way to establish the limitations of reason is through a self-critique of reason. In order for this endeavor to be consistent, a critique of reason must not employ any authority external to reason, be it political, religious, or otherwise. Only reason can be the true judge of reason. Onora O'Neill (1989) explains the importance of this point poignantly:

> Only autonomous, self-disciplining beings can act on principles that we have grounds to call principles of reason. Reason has no transcendent authority; it can only be vindicated by critique, and critique itself is at bottom no more than the practice of autonomy in thinking. Autonomy does not presuppose but rather constitutes the principles of reason and their authority. (p. 57)

Autonomy is the ultimate standard of reason and it is this standard that all rational beings ought to regard as authoritative. It is not a coincidence, then, that Kant regards autonomy as the "supreme principle of morality" (G 4:440). Autonomy is the law reason gives itself, and as such, it is the law that all rational beings ought to follow in all

practical endeavors. This principle ensures that agents are able to act without interfering with the free choice of other agents.

Knowing what autonomy is and being aware of its importance, however, is not enough. What is required is that we actually incorporate the moral law into our actions so far as possible. On a personal and practical level, this is what ought to happen with conscience; we ought to compare all our dispositions with the moral law and figure out the correct action or judgment accordingly. This requires engaging in a self-critique and finding out what we hold to be true with respect to a moral situation, without recourse to any authority except the moral law. Incorporating the moral law into our maxim and acting accordingly leads the way to regulating our conduct rationally. Typically our inclinations lead us to act in ways contrary to the moral law, but precisely by incorporating the moral law into our maxims we can govern our inclinations.

Kant states that the task of the critique is never finished, and that his own project is always open to improvement or criticism (A838/B866, A855/B883). To think otherwise would be a contradiction with the whole idea of a critique. This motto could also be applied to conscience; we do not learn conscience (in the sense of a set of particular rules or precepts which we memorize and use on every occasion), we learn to be conscientious (that is, we try to improve our capacity for judgment and strive to do what is rationally justified in each specific case). Concurrently, the work of conscience is also never finished, with each new moral case we need to consider the situation with utmost diligence and conscientiousness.

As we have seen before, according to Kant, conscience prompts us to act as if we were being judged by another person, or God. In order for this to be possible, we must be able to perceive ourselves from the point of view of another person and imagine how he/she would judge us from a moral perspective. In doing so, the principles we employ are the principles that we are morally committed to; principles that are integral in our moral constitution.

The fact that conscience presupposes moral judgment already implies that conscientiousness also presupposes a possible community in which I am judged by the maxims I adopt. The reason for this is that the material or objective rightness of actions require a third-person view of the situation. If we could not incorporate this view into our maxims, we would not be able to assess the moral import of our actions.

This could also be the reason why Kant stresses the importance of different standpoints of thought in the *Critique of Judgment*: "1. To think for oneself; 2. To think in the position of everyone else; 3. Always to think in accord with oneself" (CJ 5:294). Surely we must rely on our own reason for reaching a verdict-we must not accept a point of view without reflection. However, we must also check these verdicts and measure them from different points of view, while reflecting upon them to see if they are consistent in order that they be sufficiently intelligible. We must attempt to do our best, given our capacities and limitations, and where we judge ourselves to be insufficient we should seek the consult of our peers.

Kant relates these principles to the three higher faculties of cognition: the first principle refers to understanding, the second to the power of judgment and the third to

the faculty of reason (KU 5:294). Thinking in the position of everyone else is related to the power of judgment because this involves assessing our selves from the point of view of another, and hence it is inherently reflexive. The third rule is related to reason because it is this faculty which seeks the systematic unity of cognitions as well as actions.

The three standpoints mentioned above are also inherent in the famous categorical imperative. What usually gets lost in the discussion of the CI is that all three formulations already imply a possible community. As Onora O'Neill (1989) puts it; "the demand for universalizability is a demand that the deepest principles of our lives not preclude the possibility of community" (p. 156). Reflexive judgments, then, already imply a point of view which is not merely first-personal but takes into account a possible community.

What is of vital importance in order to engage in original, other-regarding and consistent thought and action is a proper moral education. It is to this issue that I now turn.

5.2 Methods in moral education

Moral and rational principles by themselves will not suffice for a moral education. What is needed is some understanding of the kind of being that is to be educated. In Kant's view, we require a moral anthropology for the application of these principles to human agency.

Kant defines the entirety of practical anthropology as an investigation into the vocation and character of the human being. This vocation can be satisfied by going through three educative stages of development: cultivation, civilization and moralization (AP 7:325). The cultivation of the human being contains a negative and a positive aspect; the negative aspect is discipline, which is the "taming of savagery", that is, learning to control our inclinations and the positive aspect is acquiring "skillfulness" to pursue the ends that we set forth for ourselves (P 9:449-50). This positive aspect includes the development and correct usage of ones cognitive, affective and physical abilities (P 9:475). Civilization concerns prudential reasoning (hypothetical imperatives) and finally moralization is the stage in which the individual is to attain autonomy. All three stages require education.

What Kant ultimately means by the vocation of human beings corresponds to the highest good. As I have discussed in the previous chapter, the highest good is an important concept for Kant, but it faces certain problems if it is taken so far as to include the attainment of happiness as constitutive of it. In the domain of moral education, however, I believe it can play a significant role.

This is possible because the highest good can provide a master plan for education, especially with respect to the moralization of the student. However, there is a sense in which the entirety of the human race remains a student of moralization. Accordingly, all our practical endeavors and even our speculative undertakings ought to respond to and be guided by this vocation of human beings. In this sense, I take the highest good only as a normative ideal and not as something that is bound to come about. The best we can do

is strive to be worthy of happiness and this is what moral education should aim to foster in children.

Our sensible nature is once again at issue here. Human beings inevitably follow their inclinations out of self-love. What is needed, then, is the proper regulation of inclinations so as to comply with the demands of morality. Inclinations are not to be completely eradicated, since they can be helpful, especially in the early stages of development as they can serve as a "surrogate for reason" until we reach moral maturity (AP 7:253, see also LA 25:1343).

As we have seen, moral education requires discipline, instruction and guidance. Broadly speaking, there can be two kinds of discipline for Kant. In the *Lectures on Pedagogy*, we see that discipline "prevents the human being from deviating by means of his animal impulses from his destiny: humanity" (P 9:442). The aim of this kind of discipline is to teach the student not to pursue whatever s/he happens to desire, since the result might be the violation of a moral duty. This discipline can be called self-constraint.

The other kind of discipline can be called intellectual. We can find a detailed treatment of this kind of discipline in the *Critique of Pure Reason*. Intellectual discipline aims to teach the student to be critical of others and of oneself in matters of theoretical dispute. The first step in this kind of discipline is a "thorough instruction in the critique of pure reason" (A755/B783). In this way, the student learns the limits and the basic principles of rational thought which can then be used to assess and criticize any and all claims made in theoretical disputes.

In the theoretical domain, the critique of reason has only a negative task, that is to prevent the student from uncritical/dogmatic thought (A795/ B823). The critique, then, is meant to teach us how to think for ourselves in matters of knowledge, without relying on any dogmatic authority, but also by taking into account perspectives other than our own. As always, theoretical investigation has to be a joint and harmonious effort with our peers. In the practical domain, we may think of a canon of pure reason, but here there should not be any real dispute with regards to the basic principles that ought to guide our conduct, since the moral law is universal and necessary (A750/ B778)

Improvement in any domain requires new ideas and the assessment of old ones. Kant's insistence on freedom of thought and speech (and how we cannot have one without the other) is reflected in his ideas on education. For instance, according to him, all students should be exposed to any and all kinds of ideas, even ones deemed unorthodox or even dangerous. A comprehensive education must supply the student with the right tools to engage with such ideas, which is another reason why all academic students must be thoroughly instructed in a critique of pure reason (A755/B783). This will enable pupils to get a sense of their cognitive limitations and capacities.

By engaging in this critique, the students will come to have a grasp of the limits of human cognition, as well as the principles and elements that are employed therein. The idea is that, after this instruction, the students can develop and devise argumentative strategies on their own in speculative disputes, rather than relying on merely dogmatic assertions that they have adopted from elsewhere. In other words, the students will be able to assess, criticize, accept or dismiss these ideas to their own credit, that is, freely.

In this endeavor, the teacher and the students should model their discussion upon the agreement of free citizens, so that each participant has an equal voice and carries equal weight (A739/ B767).

The best example of this kind of discussion can be found in Socratic dialogues. Indeed, Kant himself recommends the Socratic method after a certain amount of cultivation and he even alludes to the role of the teacher as the "midwife of the pupil's thoughts" (MS 6:478 and P 9:477). A characteristic feature of Socratic dialogues is that Socrates usually challenges a dogmatic assertion by insisting on a justification. His inquiry results in the interlocutor realizing that he has spoken too confidently. In other words, Socrates reminds the interlocutors of their cognitive limitations, without necessarily asserting his own view as the sole truth.

On the other hand, Socrates also encourages his peers to think for themselves and reach conclusions by using their own reason. A classic example is the dialogue entitled *Meno*. As is well-known, Socrates there attempts to teach the pupil a basic geometrical principle by appealing solely to his reason. It is only by learning to use one's own capacities that one can learn to produce original ideas and the practice of spontaneous judgment.

However, one can not begin instruction in the Socratic method, since the pupil first has to learn how to enter into a dialogue with the teacher: this requires learning to ask the right kind of questions in the right context. Therefore, the first method to be employed is the catechistic method, which aims for the cultivation of the pupil's basic cognitive skills such as memory and understanding (MS 6:478).

One reason that Kant puts forward in favor the catechistic method is that it leads the student to enjoy his/her cognitive powers and hence they could feel the progress of their faculty of judgment. In a catechistic exercise, figures and events from ancient history can be discussed with regards to their moral import, thereby activating the pupil's moral appraisal (KpV 5:154). The extension of our cognitive powers and coming across an example which relays a correct moral order can give us a kind of satisfaction which ought to be preserved and cultivated (KpV 5:160).

. However, disputes may arise with regard to the morally salient features of practical examples. In the *Metaphysics of Morals*, Kant provides several "casuistical questions" in which the teacher and the student have to work together in order to determine the right course of action (MS 6:423, *cf.* 6:426, 6:428).

One such question concerns self-sacrifice. Even though we have a strict duty against suicide (G 4:429), there might be cases in which one might be obligated to sacrifice him/herself for what s/he takes to be a greater cause. It is interesting that Kant usually leaves these questions open for discussion. The reason for this is that there might not be an easy answer, but various factors have to be accounted for in order to reach a sound conclusion. This is the kind of practice that Kant recommends for the cultivation of one's reason, rather than the memorization or repetition of the core principles of morality. As is evident from this practice, moral questions are not settled only by appealing to moral principles, but rather they require the recognition of moral features of the situation.

The issue of self-sacrifice comes to the fore in the *Critique of Practical Reason* as well. In that context, Kant recommends an example of self-sacrifice in order to show the real value of virtue.[39] According to him, the young listener "will be raised step by step from mere approval to admiration, from that to amazement, and finally to the greatest veneration and a lively wish that he himself could be such a man" (KpV 5:156). These feelings of admiration and veneration can have direct impact in the formation of the pupil's moral character, as s/he can begin to discern what is of the utmost value. The examples discussed in the catechistic method, however, should not be romantic heroes, but rather people committed to doing the right thing because it is the right thing to do (KpV 5:157).

The first step in the instruction of the student has to be dogmatic, that is, the student must first become familiar with a principle or proposition. The second step involves an attempt to refute or weaken this proposition via skeptical questioning, so the student gains a habit of criticism and does not get used to ready-made knowledge. However, stopping at this stage would not be satisfactory, and hence the student is urged to return to the basic principles of reason in order to determine whether s/he has the required capacity in order to correctly reach a conclusion. This final stage pertains to mature judgment (A761/ B 789).

Reaching mature judgment is one of the primary goals of education as it pertains to the formation of character. In Kant's conception of character we come across a three-fold distinction: a) natural aptitude, b) temperament and c) way of thinking. The first

[39] The Book of Job, discussed in the previous chapter, could be construed as such an example.

two aspects of character refer to what can be made from a human being, while the last aspect refers to what s/he can make from herself. A special emphasis is given to this third notion of *Denkungsart,* as this is what counts as the "moral predisposition" (AP 7:285).

Having this third aspect of character amounts to "that property of the will by which the subject binds himself to definite practical principles that he has prescribed to himself irrevocably by his own reason" (AP 7:292 and P 9:481). The way of thinking of a person, then, is the specific way in which an agent forms his/her own maxims and practical judgments. A person who lacks this predisposition can not count as a moral agent at all, but can only be an imitator of other people (AP 7:293).

This is why the moral education of individuals can not only consist of memorization and imitation, but must rather aim to foster the cognitive capacities that enable individuals to form their own way of thinking. Some commentators have found this puzzling and claimed that Kant's aims for education are in conflict with his ideas of autonomy.[40] If thinking originally and "giving the law to ourselves" is necessary for moral imputability, how could any form of instruction be helpful to achieving autonomy?

Kant was well aware of this difficulty. Moral education can not automatically install a character into a person's mind; there has to be an effort made by the individual. On the other hand, there have to be certain limitations in place. As Kant puts explains:

[40] See Kuehn 2006 & Louden 2016

> How do I cultivate freedom under constraint? I shall accustom my pupil to tolerate a constraint of his freedom and I shall at the same time lead him to make good use of his freedom. Without this everything is mere mechanism. (P 9:453)

The mechanistic training of human beings can never yield a moral agent (P 9:450).Again, the central feature of moral education is that the pupils learn not only moral rules, but exercise their own cognitive powers to reach morally sound conclusions. This is in line with Rousseau's ideas about the education of Emile (he shall have no leading-strings). In any kind of activity, the child should learn to participate through his/her own powers, without an overbearing authority helping out at every turn. The capacities of the students must be allowed to develop naturally, with a modicum amount of guidance.

The development of character goes hand in hand with the development of conscience. From an educational perspective, Kant regards conscience as providing an internal constraint. This constraint, Kant contends, would be the strongest if it was properly cultivated, since it belongs to the person. The proper cultivation of conscience is leads to moral perfection which is the "purpose of [one's] vocation" (LA 25:693).

In the anthropology lectures, we can see the link between morality, conscience and religion once again. As Kant puts it "This [self-constraint] cannot, however, be achieved without religion, but religion cannot have any effect without morality; hence religion aims at the highest perfection of human beings" (LA 25.695). Religion, or the recognition of our duties as divine commands discussed in the previous chapter, furnishes conscience (as well as our duties) with a certain kind of authority. However, as I have shown, the command of this authority is nothing else than having the right

disposition with regard to our duties. We cannot be mistaken with regards to our own disposition toward what morality requires.

The key point here is that our maturation has to be our own doing. This is a point that Kant stresses throughout his writings, as it relates to "thinking for oneself". This is the only way in which the spontaneity of our cognitive powers can properly be nurtured. We can not imitate spontaneity, we can only come to practice it. This was one of the ideals of the Philantropin institute, headed by Basedow. As is well-known, Kant was fully supportive of this institution and its programme.[41]

Kant's admiration for the experimental school Philanthropin, located in Dessau and founded by Johann Bernhard Basedow, can be seen in a series of short essays dated in the late 1770s as well as in his anthropology lectures. In these essays, Kant stresses the importance of education in accordance with the natural predisposition of the human beings. According to him, other educational institutes have failed because they attempted to go against nature. For this reason, they are spoiled from the start, and what is needed is not a gradual reform, but a revolution in educational practices (PH 2:449). This revolution signifies a break from a variety of educational practices which have relied extensively on memorization and recitation.

From Kant's point of view, the Philanthropin would not only educate young students, but would also give rise to well-instructed teachers in accordance with a "complete plan". Kant deems this institute as the "greatest phenomenon which has appeared in this century for the improvement of the perfection of humanity" (LA

[41] See Louden 2012 for an extended discussion of this topic.

25 :723). In many respects, the Philanthropin Institute was the best model for education available in Kant's times so it can be helpful to see some aspects of this institute and why Kant found it to be so important.

One important principle is that this institute sought to employ is that "the child cognize everything based on reasons and that it act from principles" (LA 25 :725). This is the exact opposite of a mechanistic training of children based on imitation and memorization. What this procedure aims at is the cultivation of mature judgment as originating from the student with minimal reliance on external authority.

Another important aspect of the Philanthropin institute was that it was open to children from all sects and denominations, and hence it was a cosmopolitan institute. As Jürgen Overhoff (2010) explains "the philanthropic curriculum's most important aim was to convey to the pupils those elements of their religion which connected them, not those which divided them" .[42] This is perfectly in line with Kant's account of a natural religion, and his cosmopolitan view of conscience discussed in the previous chapter. I believe Basedow's ideas were influential in Kant's general approach to religion and pedagogy.[43]

It is also interesting that Kant favored an experimental approach to education, the reason being that we can not settle one core set of educative practices for all posterity (P 9:451). Each new era faces ever new problems and requires a different kind of society (and individual) to solve them. We can not expect the following generation to strictly

[42] This quote is taken from the Dictionary of 18th Century German Philosophers under the heading Johann Bernhard Basedow.
[43] See Cavallar (2014) for the links between Basedow's and Kant's cosmopolitan approach.

follow our educative practices, since these are designed for our specific society at this specific time.

This experimental and radical approach to education might seem strange, given Kant's insistence on the universality of the moral law. After all, Kant would not have suggested that the moral law was to be altered and overthrown in the next generation. Concurrently, something similar must be the case for moral education. Even though the moral law is strict, the way in which we incorporate morality into our lives is bound to change under different times, different places and different societies. What brings these disparate engagements together is a grand plan that ought to be the aim for all humankind: the highest good.

CHAPTER 6

CONCLUSION

In this dissertation I have tried to explicate Kant's conception of conscience and how it relates to issues of moral agency, motivation, rational judgment, feeling, religion and education. I believe I was able to establish some central points about this concept and what role it plays in Kant's critical philosophy.

In my introductory chapter I have tried to put forward some common sense notions related to the concept of conscience, as well as some historical information regarding its inception and propagation. Martin Luther King's idea of a tension which gives rise to moral improvement has been incredibly inspirational for this dissertation. I believe these remarks prepared the ground for a discussion of conscience in general.

In the second chapter, I attempted to explain some of the core features of Kant's critical philosophy, especially those aspects which would help me with my discussion. Among these concepts were the moral law, freedom, reflexive judgment and moral feelings. I have also indicated that I am working within an interpretative framework in which Transcendental Idealism is not particularly central.

I believe my presentation throughout this dissertation has not needed to invoke any tenet of Transcendental Idealism, and yet, Kant's thoughts were able to speak for themselves without it. Discussions regarding Transcendental Idealism certainly have merit in themselves, but it was astonishing to see how much of Kant's moral philosophy

can be presented without it.[44] What is significant is that a clear and cogent moral philosophy can be extracted from Kant's texts without getting into the discussions regarding Transcendental Idealism.

In the next chapter I investigated the relation between conscience and religion as Kant understood it. To the best of my knowledge, none of the accounts in secondary literature has discussed Kant's interpretation of the *Book of Job* in connection with conscience. Also, I have not come across any interpretation which discusses the political implication of Kant's insistence on the infallibility of conscience. Kant was interested in biblical exegesis, for primarily political reasons as well, and I believe we can have a better grasp of Kant's understanding of religion by examining how he handles passages from both the Old and the New Testament (I do not know whether Kant wrote about any other sacred text in detail).

In the fifth chapter I have presented some connections between conscience, the critical method and moral education. I have tried to emphasize how Kant's moral and rational principles played a role in his account of education. I believe the advice Kant gives about Socratic dialogue is still relevant today and it ought to be implemented in the education of children.

I have tried to present Kant's thoughts as charitable as I can. In doing so I have disregarded some texts in which Kant was not on a par with his own moral standards. There has been a rise in recent literature about aspects of Kant's thought which are undoubtedly racist and sexist. In disregarding these texts I am neither trying to cover

[44] I am especially grateful to Kenneth Westphal for showing me how this is possible and why this is significant.

them up, nor attempting apologize for them. Thankfully, we need not treat all of Kant's writings wholesale, we can take what is valuable and leave out what is harmful. I have tried to present what I thought to be the best aspects of Kant's thought. In a manner of speaking, I have tried to present not the "letter" of Kant's philosophy, but its spirit.

What has led me to write this dissertation was ultimately my conviction that Kant's moral principles are ultimately correct, and therefore, they ought to be upheld and propagated. I hope that this dissertation will appeal not only to Kant scholars, but to any student of moral philosophy. I sincerely believe that Kant's insights about the value and nature of moral principles is still relevant today. What is more, his explorations in pedagogy and moral anthropology, while outdated, can still provide a model or aim for our endeavors in these areas.

The main reason for this is that, in all of his writings, Kant prioritizes morality, which he conceives as universal and rooted in the value of free and responsible thought and action. In doing so, he is not disregarding our sensible nature as harmful, but is rather urging us to develop ourselves in ways that are conducive to morality. This requires nurturing those aspects of our sensible nature which can be helpful in promoting morally correct action, such as conscience, benevolence and self-respect. These points are part of the reason why it is misleading to label Kantian ethics simply as "deontology". The necessity of the moral law is only one aspect of morality; according to Kant, our relation to the law as semi-rational agents is truly what determines our moral worth. The law by itself means nothing if there are no agents to implement it in the conduct of their lives.

Working out this relation between moral principles and our form of agency requires a moral anthropology which incorporates such features as religious beliefs, affective dispositions, taste, conscientiousness, moral education and so on. My dissertation has been an attempt to write out a small part of this moral anthropology as it relates to the kinds of attitudes and dispositions that we ought to have with regard to the demands of morality.

As I have shown, Kant believes that conscience, being "natural" and "incorporated into our being" is the condition for any duty to be followed in our lives. A comprehensive moral theory, therefore, can not consist of the identification and justification of objective moral norms and laws, but it must also provide an account of how we, as human beings, are supposed to act in a "morally worthy" manner. In order for the measuring-rod to be useful, it needs something to measure. Ultimately we must learn to measure our own conduct (and each others actions) in accordance with the moral law. This can only be done by a kind of education which teaches the individual not only the letter of the law, but also its spirit: a disposition of respect and sincerity toward ourselves and each other. We owe this respect and sincerity because we are able to discern what the moral law demands of us. We are rarely taught why the rights of humanity, sincerity, justice or education matters. Even if we realize the importance of rights, we rarely acknowledge that all rights ought to be complemented with corresponding duties. In this day and age, we are accustomed to demand our rights (and rightfully so, since they are regularly violated), but we do not seem to be prepared to talk about our duties or whether we are worthy of the happiness that we inevitably seek. Rampant individualism has become the law of the land. Typically we are taught the

content of the law, so that we can avoid committing crimes, just as we are educated so that we can find jobs. Things were no different two hundred years ago, as Kant was complaining of the very same problem:

> One principle of the art of education, which particularly those men who are educational planners should have before their eyes, is this: children should be educated not only with regard to the present but rather for a better condition of the human species that might be possible in the future; that is, in a manner appropriate to the idea of humanity and its complete vocation. This principle is of great importance. Parents usually educate their children merely so that they fit in with the present world, however corrupt it may be. However, they ought to educate them better, so that a future, better condition may thereby be brought forth. (P 9:447)

I hope that I will be able to direct my efforts in academia precisely in this direction. I also hope that this dissertation has been successful in working out the plausibility and urgency of this idea. If this work is able to impart some notion as to the importance of conscience, the effort was well worth it. If it further facilitates a plausible view of Kant's theory of morality in general, all the better.

First and foremost, morality demands maturity. How is this maturation going to come about? Recognizing out duties as divine commands, that is, showing the utmost respect toward the moral law is certainly a part of it. So is the cultivation of our rational as well as affective capacities (understanding, judgment, volition and also moral feeling, conscience, self-respect, love). The appreciation of beauty and compassion for all living beings could also be conducive to the maturation of the agent.

Given the variety of ways in which the demands of morality supported, it is strange that Kant is regarded as a strict German, who leaves no wiggle-room for any

personal peculiarities and demands universal obedience. In a similar vein, various formulations of the moral law have been found too abstract and hence almost irrelevant for our everyday practices. I do not see how these views could intelligibly be argued once we take a look at the variety of topics that Kant has written about. What has been the most rewarding aspect of writing this dissertation is to realize that a) Kant wrote about almost all aspects of human life and b) almost all of his texts relate back to the supremacy of morality. Morality is at the root of perpetual peace, cosmopolitanism, our sense of the sublime, our view of nature, our understanding of freedom, religion, education, our affective nature, our social relations, our political institutions and our cognitive faculties. None of these topics remain purely abstract in Kant's critical philosophy. This variety provides both richness and depth, but also focus and consistency, since all of them ultimately return to the majesty of the moral law. This oeuvre could only have been the product of an extraordinary mind and I am honored, as well as lucky, to have been able to study it closely.

I take the most significant aspect of this dissertation to be the explanation of Kant's conception of conscience in a way that supports a broadly cosmopolitan agenda. Unfortunately, our day and age is still plagued religious strife in almost all continents. We are urgently in need of a more tolerant and respectful political climate, both domestically and internationally. I am aware that this sounds extremely naive. The age we live in does not have time or patience to sift through Kant's ideas on cosmopolitanism. Those few that do can rarely rise to a position of power from which to effect meaningful change. We are, then, stuck with an insurmountable separation between mature judgment and executive power.

Has not this been the plague of humankind? Has there not been an inverse proportion between virtue and power? For the student of philosophy, this situation is not news. Socrates saw this. So did Kant. And Marx. And Arendt. That we are still struggling with the same issues is devastating. To know that these thinkers also struggled, on the other hand, could give us a glimmer of hope. Not only did they struggle, they wrote of ways to improve our condition. These texts were taken to heart by some and propagated sincerely and it is thanks to their effort that we are able to discuss, assess, share and preserve these ideas. We ought to do the same, and we are doing it, even though without much of an audience. This does not matter. We do not have much of a choice, as it is both our right and our duty to do all we can to inculcate a sense of justice and virtue in our prospective students, in light of these most valuable texts.

If this dissertation is what I leave for posterity, then, may it speak of the importance of conscientiousness and the worthiness to be happy.

REFERENCES

Allison, H. (1990). *Kant's theory of freedom*, Cambridge: Cambridge University Press.

Aquinas, T. (1947). *Summa theologica*, Fathers of the English Dominican Province (Tr.), Benziger Bros, Retrieved from: http://dhspriory.org/thomas/summa/FP/FP079.html#FPQ79A13THEP1

Bacin, S. (2013). The perfect duty to oneself merely as a moral being (TL 6:428-437). In O. Sensen, J. Timmermann & A. Trampota (Eds.), *Kant's Tugendlehre: A Comprehensive Commentary* (pp. 245-268), Berlin/New York: Walter de Gruyter.

Baxley, A. M. (2010). *Kant's theory of virtue: The value of autocracy*, Cambridge: Cambridge University Press.

Cavallar, G. (2014). Sources of Kant's cosmopolitanism: Basedow, Rousseau and cosmopolitan education, *Studies in Philosophical Education*, *33* (4), 369-389.

Coogan, M. D. (2010). *The new Oxford annotated Bible: New revised standard version with the Apocrypha*, Oxford: Oxford University Press.

Denis, L. & Sensen, O., (2015). Introduction. In L Denis & O. Sensen (Eds.), *Kant's lectures on ethics: A critical guide* (pp 1-12), Cambridge: Cambridge University Press.

Esser, A. (2013). The inner court of conscience, moral self-knowledge, and the proper object of duty (TL 437-444). In O. Sensen, J. Timmermann & A. Trampota (Eds.), *Kant's Tugendlehre: A comprehensive commentary* (pp. 269-291), Berlin/New York: Walter de Gruyter.

Formosa, P. (2011). A life without affects and passions: Kant on the duty of apathy, *Parrhesia*, *13*, 96-111.

Green, T. (1999). *Voices: The educational formation of conscience*, Indiana: University of Notre Dame Press.

Guyer, P. (1997). *Kant and the experience of freedom*, Cambridge: Cambridge University Press.

Guyer, P: (2000). *Kant on freedom, law and happiness*, Cambridge: Cambridge University Press.

Guyer, P. (2003). Kant's principles of reflecting judgment. In P. Guyer (Ed.), *Kant's Critique of the Power of Judgment: Critical essays* (pp. 1-62), Lanham: Rowman & Littlefield.

Guyer, P. (2010). Moral feelings in the Metaphysics of Morals. In L. Denis (Ed.), *Kant's Metaphysics of Morals: A critical guide* (pp. 130-151), Cambridge: Cambridge University Press.

Hare, J. (2011). Ethics and religion: Two Kantian arguments, *Philosophical Investigations, 34* (2), 151-168.

Herman, B. (1993). The practice of moral judgment, *The practice of moral judgment*, London: Harvard University Press, 73-93.

Herman, B. (1998). Training to autonomy: Kant and the question of moral education. In A. O. Rorty (Ed.), *Philosophers on education* (pp. 254-271), London: Routledge.

Herman, B. (2007). *Moral literacy*, Cambridge, MA: Harvard University Press.

Hill, T. E. (2002). *Human welfare and moral worth: Kantian perspectives*, Oxford: Oxford University Press.

Kain, P. (2005). Interpreting Kant's theory of divine commands, *Kantian Review, 9* (1) ,128-149.

Kant, I. (1963). *Lectures on ethics* (L. Infield, Trans.), New York: Harper & Row Publishers.

Kant, I. (1996). *Practical philosophy*, M. J. Gregor (Trans. & Ed.), Cambridge: Cambridge University Press.

Kant, I. (1996)b. *Religion and rational theology* (A. Wood & G. Giovanni,Trans. & Eds.), Cambridge: Cambridge University Press.

Kant, I. (1997). *Lectures on ethics* (P. Heath, Trans.), J. B. Schneewind & P. Heath (Eds.), Cambridge: Cambridge University Press.

Kant, I. (1998). *The critique of pure reason* (P. Guyer & A. Wood, Trans.) Cambridge: Cambridge University Press. (Original work published 1781).

Kant, I. (2000). *The critique of the power of judgment* (P. Guyer & E. Matthews Trans.), Cambridge, Cambridge University Press. (Original work published 1790).

Kant, I. (2005). *Notes and fragments* (P. Guyer, Eds. & Trans.), Cambridge: Cambridge University Press

Kant, I. (2007). *Anthropology, history and education* (M. Gregor, P. Guyer, R. Louden, H. Wilson, A. Wood, G. Zöller & A. Zweig, Trans.), G. Zöller & R. Louden, (Eds.), Cambridge: Cambridge University Press

Kant, I. (2012). *Lectures on anthropology* (R. Clewis, R. Louden, F. Munzel, A. Wood, Trans.), A. Wood & R. Louden (Eds.), Cambridge: Cambridge University Press

Kazim, E. (2017). *Kant on conscience: A unified approach To moral self-consciousness*, Leiden, Netherlands: Brill.

King, M. L. (1963). Letter From Birmingham Jail, Retrieved from: https://www.africa.upenn.edu/Articles_Gen/Letter_Birmingham.html

La Rocca, C. (2016). Kant and the problem of conscience, *Contemporary Studies in Kantian Philosophy, 1*, 65-79.

Louden, R. (2011). *Kant's human being*, Oxford: Oxford University Press.

Louden, R. (2016). Total transformation: Why Kant did not give up on education, *Kantian Review, 21* (3), 393-413.

Makkreel, R. (2002). Reflective judgment and the problem of assessing virtue in Kant, *The Journal of Value Inquiry, 36* (2-3), 205-220.

Moyar, D. (2008). Unstable autonomy: Conscience and judgment in Kant's moral philosophy, *Journal of Moral Philosophy, 5* (3), 327-360.

Paton, H.J. (1960). Conscience and Kant, *Kant Studien, 70* (3), 239-251

Pascal, B. (1995). XLV-680, *Pensees and other writings*, H. Levi (Tr.), Oxford: Oxford University Press.

Plato (1997). *The complete works of Plato*, J. Cooper (Ed.), Indianapolis, Hackett Publishing

O'Neill, O. (1985). Consistency in action. In N. Potter & M. Timmons (Eds.), *Morality and universality* (159-186), Dordrecht, Holland: Reidel Publishing Company

O'Neill. O. (1989). *Constructions of reason: Explorations of Kant's practical philosophy*, Cambridge: Cambridge University Press.

O'Neill, O. (1996). Kant on reason and religion, *The Tanner lectures on human values*, Retrieved from: https://tannerlectures.utah.edu/_documents/a-to-z/o/oneill97.pdf, 269-308.

Sherman, N. (2014). The place of emotions in Kantian morality. In A. Cohen (Ed.), *Kant on Emotion and Value* (pp. 11-32), Hampshire: Palgrave Macmillan.

Sorabji, R. (2014). *Moral conscience through the ages: Fifth century BCE to the present*, Oxford: Oxford University Press.

Sorensen, K. (2002). Kant's taxonomy of emotions, *Kantian Review, 6* (1), 109-128.

Sticker, M. (2016). When the reflective watch-dog barks: Conscience and self-deception in Kant, *The Journal of Value Inquiry*, *51* (1), 85-104.

Strohm, P. (2011). *Conscience: A very short introduction*, Oxford: Oxford University Press.

Timmons, M. (2002). Motive and rightness in Kant's ethical system. In M. Timmons (Ed.), *Kant's Metaphysics of Morals: Interpretative essays* (pp. 255-288), Oxford: Oxford University Press.

Vujosevic, M. (2014). The judge in the mirror: Kant on conscience, *Kantian Review, 19* (3), 449-474.

Westphal, K. (2004). *Kant's transcendental proof of realism*, Cambridge: Cambridge University Press.

Westphal, K. (2010), Practical reason: Categorical imperative, maxims, laws. In W. Dudley & K. Engelhard (Eds.), *Kant: Key concepts* (pp. 103-119), London: Acumen.

Westphal, K. (2011), Kant's moral constructivism & rational judgment. In S. Baiasu, S. Pihlström & H. Williams (Eds.), *Politics and metaphysics in Kant* (pp. 28-46), Cardiff: Wales University Press.

Westphal, K. (2017a-forthcoming), Universal moral principles & mother wit. In M. Kaufmann, J. Thompson, M. Massa & S. Brandt (Eds.), *Regelfolgen ,regelschaffen, regeländern – die Herausforderung für auto-nomie und universalismus durch Ludwig Wittgenstein, Martin Heidegger und Carl Schmitt*, Frankfurt am Main: Peter Lang

Westphal, K. (2017), How Kant justifies freedom of agency without transcendental idealism, *European Journal of Philosophy*, 25, DOI: 10.1111/ejop.12264

Wood, A. (2008). Conscience, *Kantian ethics*, Cambridge: Cambridge University Press, 182-192.

Wood, A. (2014). The evil in human nature. In G. Michalson (Ed.), *Kant's Religion within the Boundaries of Mere Reason: A critical guide* (pp. 31-57), Cambridge: Cambridge University Press.