A Study on Conscience: The Content and Function

Nalan Sarac

Follow this and additional works at: https://dsc.duq.edu/etd

Recommended Citation

This One-year Embargo is brought to you for free and open access by Duquesne Scholarship Collection. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Duquesne Scholarship Collection. For more information, please contact phillipsj@duq.edu.
A STUDY ON CONSCIENCE: THE CONTENT AND FUNCTION

A Dissertation
Submitted to the McAnulty College & Graduate School of Liberal Arts

Duquesne University

In partial fulfillment of the requirements for
the degree of Doctor of Philosophy

By
Nalan Saraç

August 2016
Copyright by

Nalan Saraç

2016
A STUDY ON CONSCIENCE: THE CONTENT AND FUNCTION

By

Nalan Saraç

Approved July 7, 2016

Ronald Polansky
Professor of Philosophy
(Committee Chair)

Thérèse Bonin
Associate Professor of Philosophy
(Committee Member)

Jennifer Ann Bates
Professor of Philosophy
(Committee Member)

James Swindal, Dean
The McAnulty College and Graduate School of Liberal Arts
Professor of Philosophy

Ronald Polansky
Chair, Department of Philosophy
Professor of Philosophy
ABSTRACT

A STUDY ON CONSCIENCE: THE CONTENT AND FUNCTION

By
Nalan Saraç
August 2016

Dissertation supervised by Ronald Polansky

The aim of this study is to contribute to the understanding of conscience by critically examining turning points of the traditionalist approach that conceptualizes conscience primarily as a cognitive capacity. The basic assumption of this approach is that there is a built-in mechanism in human beings that enables us to judge the rightness of our actions. This mechanism is a part of reason that deals with the ethical value of the actions. Conscience judges an action with reference to a set of principles. These principles constitute the content of conscience. My study follows this track and puts forth its findings in terms of its content and function.

The study starts with the examination of the emergence of the word conscience. Once I have enough evidence that enables me to fix the meaning and use of the word I turn to Plato and Aristotle and argue that although they do not write explicitly on
conscience they establish the conceptual framework of the later studies on conscience. In the second chapter, I focus on Philip the Chancellor, Bonaventure, and Aquinas’ accounts of conscience and claim that the importance of these thinkers arises from the questions they ask regarding the relation between conscience and reason. They highlight the question that has to be answered in order to explain the mechanism of conscience. In the third chapter I focus on Kant’s understanding of conscience. He provides coherent answers to the questions about the distinction between reason and conscience. I argue that his Copernican Revolution in ethics enables him to re-secularize the concept of conscience that originally emerges as a secular one. In the fourth chapter, I consider Nietzsche’s criticism about conscience. I conclude that his criticism is not so controversial as it is thought and his main contribution to the studies on conscience is that the content of conscience may not be some necessary truths about the right thing to do but may come from contingent beliefs, which are imposed on us by the society in which we are brought up.

In the conclusion, I claim that the authority of conscience arises neither from the contents, whatever they be, nor from the ways conscience derives conclusions, but from the belief that I could not live with myself if I do something wrong, so I have to scrutinize my actions and judgments. If I find something wrong about my past action I have to admit the guilt, or if I find something wrong about my future action, I should not do it. This belief originates from one’s awareness of one’s accountability for one’s actions.
DEDICATION

For Ceren
ACKNOWLEDGMENT

I wish to give my heartfelt thanks to my dissertation director Dr. Ronald Polansky and committee members Dr. Jennifer Bates and Dr. Thérèse Bonin for helping me to develop my work. I also wish to thank my colleagues at the Koç University, Dr. Charilaos Platanakis and Dr. Eylem Özaltun and Dr. Hülya Şimga; for their comments and encouragement. And, finally, I would like to thank to my family: my parents Fahriye and Necdet Saraç, for their love and support, without their help I would not have been able to finish this study; my husband, Memduh Er, for encouraging me to start the PhD program and endure the hardship of this process, and my daughter, Ceren for her patience and inspiration.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>iv</td>
</tr>
<tr>
<td>Dedication</td>
<td>vi</td>
</tr>
<tr>
<td>Acknowledgement</td>
<td>vii</td>
</tr>
<tr>
<td>Introduction</td>
<td>x</td>
</tr>
<tr>
<td>Chapter 1: From Ancient Greece to 1st Century Europe</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Emergence of the term/phenomenon</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Plato</td>
<td>15</td>
</tr>
<tr>
<td>1.3 Aristotle</td>
<td>23</td>
</tr>
<tr>
<td>1.4 Hellenistic Period</td>
<td>31</td>
</tr>
<tr>
<td>1.5 Conclusion</td>
<td>43</td>
</tr>
<tr>
<td>Chapter 2: Medieval Understanding of Conscience</td>
<td>45</td>
</tr>
<tr>
<td>2.1 Background of the Discussion</td>
<td>45</td>
</tr>
<tr>
<td>2.2 Peter Lombard</td>
<td>48</td>
</tr>
<tr>
<td>2.3 Philip the Chancellor</td>
<td>51</td>
</tr>
<tr>
<td>2.4 Bonaventure</td>
<td>65</td>
</tr>
<tr>
<td>2.5 Thomas Aquinas</td>
<td>78</td>
</tr>
<tr>
<td>2.6 Conclusion</td>
<td>93</td>
</tr>
<tr>
<td>Chapter 3: Kant</td>
<td>96</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>96</td>
</tr>
<tr>
<td>3.2 Conscience as a Moral Feeling</td>
<td>99</td>
</tr>
<tr>
<td>3.3 Conscience as an Inner Court of Moral Judgment</td>
<td>105</td>
</tr>
<tr>
<td>3.4 On the Possibility of an Erring Conscience</td>
<td>114</td>
</tr>
<tr>
<td>3.5 Conclusion</td>
<td>119</td>
</tr>
<tr>
<td>Chapter 4: Nietzsche</td>
<td>122</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>122</td>
</tr>
</tbody>
</table>
4.2 Nietzsche and Morality .................................................. 123
4.3 Development of the Bad Conscience .................................. 125
4.4 Emergence of Conscience .................................................. 139
4.5 Conclusion ........................................................................ 142
Conclusion ............................................................................. 144
Bibliography ............................................................................ 153
Introduction

Statement of the Problem

Whenever there is a crisis involving apparent culpability people “appeal to conscience”. However, the phrase is ambiguous. First, it may mean an appeal to another person’s conscience in order to convince him to act in certain ways. Equally, it may refer to the invocation of one’s own conscience to interpret and justify one’s conduct to others. Finally, it may connote the role of conscience in debates with oneself about the right course of action, conscience being understood as a participant in the debate, a referee, or a final arbiter (Childress 1979, 315). The ambiguity of the phrase arises from the concept of conscience itself. According to Merriam-Webster it is:

1. The part of the mind that makes you aware of your actions as being either morally right or wrong.
2. a feeling that something you have done is morally wrong.
3. the sense or consciousness of the moral goodness or blameworthiness of one’s own conduct, intentions, or character together with a feeling of obligation to do right or be good.
4. a faculty, power, or principle enjoining good acts

Even in a single dictionary entry, it refers to many different entities. These include: a part of the mind, a feeling, a sense, a faculty, a power, and finally a principle. Which one is the actual referent of the term conscience? Or does it have any clear referent? Or is it merely a rhetorical concept? If it is rhetorical, what is the use of it?

---

In any case it refers to a supposed agency in us “disapproving wrong acts”; “enjoining good acts”; it judges our performance and orders what should be done in the future. Thus conscience is about the activity of some agency of the self in self-judgment and idealization. In other words, one judges oneself and models oneself in prospect in the light of one’s ideal. This understanding of conscience is the basis of various interpretive theories that can be listed under two headings: traditionalist and non-traditionalist.²

Traditionalists generally assume that there is in most of us an agency that serves as a final court of appeal in making ethical decisions. They view conscience as an inherent, God-given, or natural capacity for making ethical judgments. If one acts contrary to its instructions, the torments of guilty conscience will follow. Conscience then has a fixed content that is in tune with the “law of God” or the “natural law”. Starting with the Stoics and early Christian thinkers until the 19th century, the traditional approach was the dominant paradigm of the studies of conscience. The non-traditionalist view, inspired by the works of Nietzsche and Freud, takes conscience as a social and psychological phenomenon. According to this approach, conscience derives primarily from internalizations of social values. Since it is not natural or God-given but formed by fallible social interactions, its dictates are seen as inhibitions rather than guideposts. Thus, the first approach regards conscience as something natural, as an authority built into the structure of selfhood that guides one to stick to her true self. The second views it as an external imposition, as a form of domination that must be overcome in order for the individual to achieve-self-fulfillment.

² This is a distinction I draw on the basis of my research.
There is a track in the traditionalist approach that examines conscience primarily as a cognitive capacity.\(^3\) The aim of this study is to follow this track and put forth its findings in terms of the content and function\(^4\) of conscience in a critical way and then discuss the significance of it. The basic assumption of this approach is that there is a built-in mechanism in human beings that enables us to judge the rightness of our actions. This mechanism is a part of reason that deals with the ethical value of the actions. Conscience judges an action with reference to a set of principles. These principles constitute the content of conscience.

Early contributors to the approach assume that there is a fixed content of conscience whereas the later contributors emphasize the importance of the individual’s role in the content formation. Each individual interprets the norms and values of her age or society with the help of her own experiences and reason and comes up with an idea of how she should be. In this way, she creates an ideal self. The principles of the ideal self about the right action constitute the content of the conscience. So, both reason (mechanism) and the ideal self (content/principle) are the constituents of conscience. Nevertheless, studies on conscience have to clarify what is distinct about it by showing its difference from reason and the ideal self. In this way this project can fulfill its promise.

---

\(^3\) To limit the scope of this study I leave the thinkers, who conceptualize conscience as an emotional capacity, and focus on only the thinkers who examine conscience primarily as a cognitive capacity. There are two reasons for this choice. First, the literature on conscience as a cognitive capacity is richer than the conscience as an emotional capacity. Second, those who approach on conscience as an emotional capacity, especially Hume and J. S Mill, identify it with a sensation of pain or discomfort without paying attention to the content and function of conscience. Actually these sensations are a part of the experience of conscience but without one’s belief about right and wrong their accounts are not sufficiently comprehensive enough.

\(^4\) By function I mean both the telos of conscience and its mechanism, i.e., how it works. Depending on the context I will use function and mechanism interchangeably.
Conscience has typically been experienced as the voice of one’s ideal self, who gives an argument following a principle, as it is exemplified in the following argument:

P1) I ought to X because X is the right thing to do.

P2) Y is not an instance of X because Y violates X.

C) Therefore, Y is not the right thing to do.

Furthermore,

C1) If I have Yed already, I did something wrong and I am guilty OR,

C11) If I have not Yed yet, I ought not Y.

Obviously, P1 is based on a principle that one ought to do the right thing. For example, telling the truth is something one ought to do. Nevertheless, conscience is not equal to following the orders of an ideal self because if it were, a reference to conscience as a reason for action would preclude the possibility of personal, ethical decision. For acting conscientiously would merely be doing what one has been told to do—whether by the voice of God or the voice of an ideal self is irrelevant—rather than doing what one has decided as right.

Similarly, the mechanism of conscience draws the conclusion that an action is wrong by reasoning but it is not equal to ethical reasoning. It is the mechanism of conscience that enables reason to infer C1 and C11 from C. It is a specific use of reason and we need to explain why it works in this way. In other words, what does trigger this mechanism? Is there another part of reason that makes conscience work or do we have a disposition to evaluate all of our actions in this way or is this something we have to learn to use?
These and many other questions preoccupied the philosophers who for centuries wanted to develop an account of conscience. Until the 19th century, they also tried to clarify the content of the conscience, though it was generally assumed that the content of conscience is provided by God. Because of this assumption, they worked primarily on the mechanism of conscience. In the 19th century, the authority of God as the only content provider for conscience had been challenged and other sources had been identified for providing the content for conscience, namely, the individual, family and society. Therefore, the content of conscience became the focus of conscience studies.

An examination of the emergence of the concept conscience is crucial. A historical survey of the origin of the concept can provide a basis for our understanding of its use in the philosophical context. For this reason, it will be necessary to analyze its non-philosophical uses as well.

Although neither of them wrote explicitly on conscience, I believe it is appropriate to include Plato and Aristotle because their understanding of phenomena later connected with conscience greatly influenced their successors and served as a basis for the discussions on conscience.

Plato’s understanding of the self played a crucial role in the establishment of the content of conscience. Plato, in his dialogues often emphasized the importance of being aware of one’s own misdeeds, weaknesses, or contradictions. In fact, this awareness is the very condition of being an ethical person. Plato conceptualizes the self as a self-controlled and harmonious unity under the rule of reason. Therefore, it is necessary to know what is going on in every constituent of this unity. So, the reason can control them and put them in order in accordance with a vision of natural order. If one fails to detect
incompatible desires, beliefs, attitudes in the self and put them in order one will end up with disharmony, i.e., injustice in his soul, which is equal to being sick.

This disturbance is explained with a metaphor in the *Hippias Major* (304c-e) where Socrates is questioned and insulted by a close relative about his misdeeds and contradictions when he comes home. Coming home is a metaphor for thinking in the sense of examining ourselves in terms of our deeds, thoughts, beliefs, and values. Whenever we think, we engage in a silent dialogue with ourselves, there is a possibility of encountering a contradiction within ourselves that can distort the harmony of our soul. The fear of having a disharmonious soul and a blaming and insulting relative at home is what conscience is. It compels us to check our deeds and words in the light of reason in order to live peacefully with ourselves. Hence we can conclude that fear of knowing within oneself that one is inconsistent lies at the heart of Plato’s ethics.

Plato’s emphasis on the importance of living according to the principles of the self constituted the content of conscience. However, later thinkers replaced the principles of the self with the principles of God, starting with Paul. Hence the ideal self turned into God’s image in human beings and the dictates of God became the content of conscience.

Aristotle described the mechanism of reason in practical matters and argued that in order to lead an ethical life we have to prepare ourselves for contingencies by developing practical wisdom that guides us to do the right thing. As a capacity to differentiate right from wrong, practical reason is similar to our modern understanding of conscience. Although practical wisdom tells us what the right thing to do is, sometimes we fail to do it. According to Aristotle, the *reason* for this failure is that the incontinent person does not cultivate moral virtues, so that he cannot develop practical wisdom fully.
His account of weakness of will exemplifies such a failure. Aristotle’s examination of weakness of will/akrasia provides a paradigm to explain how we fail to do the right thing although we know the guiding principles; in other words why we fail to listen to the voice of conscience.

Aristotle’s account of practical wisdom and weakness of will played an important role in the later studies on conscience. Medieval thinkers discussed conscience in Aristotelian terms. In addition, they were influenced by Paul’s conceptualization of conscience in relation to the natural law in Rom. 2:14 – 15. Starting with Philip the Chancellor, medieval thinkers drew a distinction between synderesis and conscientia. Although there are textual grounds behind this distinction, it is also motivated by a desire to keep at least a part of conscience from failure. Synderesis, the infallible part of conscience is called the spark of conscience, consists of basic ethical principles, i.e., following the terminology of this study, it is the content of conscience. Although medieval thinkers do not doubt that the source of these principles is God, they investigate the way human beings acquire them. Conscientia is the process of drawing particular conclusions from the general principles, i.e., the mechanism. Since it relies on reason, its operations are subject to error. Following Philip the Chancellor, Bonaventure and Aquinas discussed whether synderesis (and conscientia also) is a potentiality, a disposition or an actualization; does it belong to the desiring or thinking part of the soul; can it be mistaken? Critical examination of all of these questions is necessary to decide whether the examination of conscience in terms of mechanism (conscientia) and content (synderesis) is really significant.
Kant’s account of conscience also has to be taken into consideration because although he mainly follows the same track, he re-secularizes the content of conscience and also depicts the role of the individual in the process of content formation. Like his predecessors Kant also believes that “there is a germ of goodness left in its entire purity, a germ that cannot be extirpated or corrupted” (Kant 1793, Religion. Ak. 6:45-46). In other words, there is some basic sense of goodness that is the source of feelings such as remorse. Nevertheless, Kant also asserts that man is autonomous; in the sense that he is the legislator of his own law. At the same time man is the subject of these laws. They are formulated by practical reasoning. It is the duty of conscience to check whether man follows the orders of the law. Therefore, “consciousness of an inner court in the human being (‘before which his thoughts accuse or excuse one another’) is conscience” (Kant 1797, Metaphysics ofMorals, Ak. 6:438). He describes conscience as the duty to engage in a kind of second-order reflection, judging that one has applied moral judgments properly to oneself.

Kant’s account highlights the distinction between the source of ethical knowledge and the motivation for being ethical. This distinction has to be elaborated in order to deepen our understanding of conscience because it provides insights to explain the mechanism of conscience.

The assumption that there is a built in mechanism with a fixed and/or developed content that enables human beings to judge the ethical value of their action is challenged by the thinkers of the 19th century. As a prominent figure of this movement, Nietzsche provides rich resources. In On the Genealogy of Morals, Nietzsche describes conscience as a sickness that is caused by the transition from wild life to the civilized life. The
instincts of hunting, cruelty, hostility, and destruction that belong to the wild beast had to be suppressed during civilization. For this reason, human beings turned all this toward themselves, made themselves a new wilderness to be struggled against and conquered. In so doing, an inner life and bad conscience emerged. Contrary to his predecessors, Nietzsche conceptualizes conscience as something questionable, as an impediment for human beings, which prevents them from realizing their true nature. This controversial view can give us an opportunity to check our presuppositions about human nature. For example, do human beings, as a species have a real concern for doing the good thing? If not, then conscience is nothing but a tool for controlling human behavior to maintain the social order.

This study aims to put forth a critical analysis of the accounts that conceptualize conscience in terms of its content and mechanism. All of the accounts discussed above represent turning points along these lines. In the analysis, the focus will be on their explanation of how and why conscience works. The results of this analysis should provide a basis for a discussion about the significance of conscience for us today. In this way this study will contribute to the philosophical understanding of conscience.

Literature

There are certain tracks in conscience studies, such as New Testament studies or studies of medieval texts. The earliest studies on conscience began in the 19th century. The scholars of the 19th century were interested in the literary sources of antiquity and the New Testament in order to find out the origin of conscience and its significance for religion and ethics. Martin Kähler’s Das Gewissen: Ethische Untersuchung. (1878) is one of the prominent works on conscience. He claims that the term should be viewed within
the context of general conceptions of ethics in the ancient world. He argues that changes in the social circumstances in late 5th century Athens, made people question the traditional values and hence conscience stemmed from the relativity or even the invalidity of traditional values. The individual rejected the traditional views about right and wrong because the authority of those views had been undermined. However, in the very process of rejection, he experienced himself being checked by an inner form of control. The process of internalization was a decisive factor for the genesis of conscience.

In the 20th century, scholars in the field of New Testament studies were interested in the topic and published many works. C. A. Pierce’s *Conscience in the New Testament* (1955) is the most important of them. He listed all the appearances of the *sunoida* word group between the 5th century BC and 3rd century AD and classified them according to their form and usage. He concluded that *suneidesis / conscience* is “the pain suffered by man, as man, and therefore as a creature in the order of things, when by his acts completed or initiated, he transgresses the moral limits of his nature” (p. 54). He claimed that the term belonged to the socio-cultural environment of Paul, who used it for the first time in the Christian literature, and should be considered authoritative for Christianity as a whole.

Following this tradition, Philip Bosman takes up the issue in his recent book *Conscience in Philo and Paul: A Conceptual History of the Sunoida word Group* (2003). He claims that recent developments in cognitive science on memory and the way information is stored in the brain has opened new perspectives on the nature of concepts and conceptuality. In combination with strict philology, these insights may renew our understanding of the ancient roots of conscience. In his detailed study he lays down the conceptual framework featuring the *sunoida* word group. He concludes that transgressing the accepted moral code produces various disturbing psychological reactions and the ancient Greeks conceptualized them as inner turmoil, which inevitably has detrimental and even destructive consequences for the individual. The *sunoida* word group emerged and evolved to designate this awareness and emotions.

In the second track, there are studies on conscience in the medieval texts. The most important work in this field is Lottin’s *Psychologie et morale aux XIIe et XIIIe siècles* (1948). He compiled almost all the medieval texts on conscience. Although his work is a very important resource for conscience studies, because of his scholastic style, it is not accessible for non-specialists. Timothy Potts undertakes the task of making these texts more accessible to the readers by interpreting the medieval discussion of conscience in *Conscience in Medieval Philosophy* (1980). He illuminates the context of the discussions by explaining the Aristotelian terminology and how medieval writers employed it in order make sense of conscience.

Douglas Langston’s *Conscience and Other Virtues: From Bonaventure to MacIntyre* (2011) can be listed among the followers of the second track. Although he includes Luther, Butler, Kant Freud, Ryle, and MacIntyre in his study, his main focus is
medieval conceptualization of conscience and its significance for virtue ethics. He argues that medieval discussions of conscience provide us with strong grounds for claiming the importance of conscience in virtue ethics because conscience plays a crucial role in the training of virtues.

Paul Strohm’s *Conscience: A Very Short Introduction* (2011), as its title implies, gives brief information about its use in philosophy, religion, and literature in different periods starting with Cicero and Augustine through the Middle Ages and into the Reformation. He argues that conscience is an important concept for the discussions of human rights and in contemporary politics.5

In addition to these two main tracks there are studies on conscience, which deal with one philosopher’s conceptualization of conscience, such as Allen Wood’s “Kant on Conscience”, David Jones “Freud’s Theory of Moral Conscience”, or Aaron Ridley’s “Nietzsche’s Conscience”, among many others.

Although the existing literature is illuminating for understanding how conscience is conceptualized by the philosophers in certain periods of time, almost none of the literature tries to provide a structure for its analysis. This study aims to fill this gap in the literature by bringing together the accounts of conscience by means of examining conscience in terms of its content and mechanism.

5 While I was working on this project, Richard Sorabji produced a detailed study on conscience: “Moral Conscience through the Ages: Fifth Century BCE to the Present” Nov. 2014. He attempts to answer to the question “what is moral conscience?” He analyzes philosophical, religious, and political texts to provide a clear understanding of conscience. He pays special attention to the concept freedom of conscience in religious and political contexts. His findings regarding the emergence and development of the concept in the Ancient Greece are for the most part parallel to my conclusions in the first chapter for the most part.
Methodology

To develop a critical analysis of the accounts that conceptualize conscience in terms of its content and mechanism, I critically examine the turning points in the history of conscience that contribute significantly to this approach. Although conscience is not a concept on which philosophers write extensively, it is still possible to find at least a couple of pages on conscience in almost every philosopher who is concerned about ethical matters. However, most of them dwell on the existing ideas without making significant contributions. I focus on those that broaden our understanding of conscience by developing a new conceptual framework.

The study starts with the examination of the emergence of the word conscience. I go over the *Thesaurus Linguae Graecae* search results for ‘sunoida’, which is considered to be the origin of conscience and pick out the ones that display the characteristics of the meaning of the word. Once I have enough evidence that enables me to fix the meaning and use of the word I turn to Plato’s dialogues such as *First Alcibiades, Republic, Apology, Crito, Symposium, Gorgias, and Hippias Major* and consider what is relevant in them to our topic. Although Plato does not use the word frequently, his requirements for being an ethical person constitute the conceptual framework of conscience. Then I analyze Aristotle’s understanding of practical wisdom and weakness of will that serve as the basis for the later discussions on conscience. I analyze Hellenistic philosophers’ and Paul’s use of conscience because they also influence later studies on conscience.

In the second chapter, I focus on Philip the Chancellor, Bonaventure, and Aquinas’ accounts on conscience because these thinkers develop detailed analyses of conscience with respect to its mechanism and content. In the third and fourth chapters I
work on single philosophers; Kant and Nietzsche whom I consider as the representative of a certain approach. I focus on Kant’s understanding on conscience as a representative of enlightenment. His emphasis on reason and self-knowledge enables him both critically review the findings of his predecessors and brings new insights to the concept. In the fourth chapter, I represent Nietzsche’s criticism about conscience as a challenge to the traditionalist approach by paying special attention to his *On the Genealogy of Morals*. The final chapter brings together the findings of this study. I try to determine what for us is a compelling account of the content and mechanism of conscience.
CHAPTER I

From Ancient Greece to 1st Century Europe

1.1 Emergence of the term/phenomenon

The aim of the chapter is to trace the development of the term conscience from its emergence in the 6th century BC to the 1st century AD, when its meaning is fixed and becomes a standard philosophical term. First, I focus on the word itself, then I explore the conditions that give rise to the emergence of the term and argue that only a certain kind of self can have conscience. Finally, I attempt to examine its use in the ethical context by focusing on the writings of the philosophers, who make significant contributions either by creating a framework for the concept or adding new insights to it.

The word “conscience” derives etymologically from the Latin conscientia, which looks like a verbatim translation of the Greek term suneidēsis (συνείδησις). Although there are competing viewpoints about which one precedes the other, textual evidence supports the precedence of the Greek term. In addition, since the ethical use of the term was pervasive during the Hellenistic era, it would not be unreasonable to assume that Latin writers derived their use of conscientia from Greek sources. For this reason, an investigation of the emergence of the term suneidēsis can be a good starting point for our quest for the meaning of conscience.

---

6 First appearance of suneidēsis in Greek literature is in the 5th century BC in Democritus: (DK, fr.297), whereas conscientia appears for the first time in Latin literature in the first century BC in the Rhetorica ad Herennium.
The group of words and phrases to which *suneidēsis* belongs, is found throughout Greek writings from the 6th century BC to the 7th century AD. All kinds of writers used it including philosophers, poets, tragic and comic playwrights, historians, physicians, rhetoricians, and writers of private correspondence. A. C. Pierce claims that “it is in fact an ‘everyday’ group of words expressing a commonplace idea — truly popular, and belonging rather to ‘folk-wisdom’ than to ‘popular philosophy’ — or, rather, second-hand philosophical jargon (Pierce 1955, 16-17).

The word, from which all the words and phrases in this group derived is *sunoida* (*συνοιδα*- inf. *συνειδέναι*). It is a composite verb with two parts:

I. The prepositional prefix *σύν/ξύν* is used with dative. It means “along with, in company with, together with.” In composites, it means “with, along with, together, at the same time, hence of any kind of union, connexion, or participation in a thing, and metaphorically agreement or unity.”

II. The word *oida/oĩda* is perfect indicative of the verb *eidō/εἰδω* (see, perceive). It means “to know.” It implies knowing intuitively as opposed to acquiring knowledge through reasoning (*noein/ νοεῖν*) (Marietta 1970, 177).

Hence the composite *συνοιδα* has the following meanings, depending on the grammatical construction:

1. to share in knowledge, be cognizant of a thing, be privy to it,

---


2. ἑαυτῷ συναιδέναι τι to be conscious of a thing

   a. in nom., εἱ. ἐμαυτῷ οὐδ’ ὁτιοῦν σοφὸς ὁν Plat.; without the reflex. Pron. to be conscious that, ξύνοισθά γ᾽ εἰς ἐμ’ οὐκ εὐδοκοι ὁν Eur.

   b. in dat., εἱ. ἐμαυτῷ οὐδὲν ἐπισταμένῳ I am conscious that I know nothing, Plat.

   c. in acc., ξόνοιδ’ Ὀρέστην σε ἐκπαγλουμένην I know well that thou admirest him, Aesch.

3. absol. ξυναιδός, an accomplice, εἱ. τις Thuc.; also, ὁ εἱ. τινι Thuc.

   b. neut. τὸ συναιδὸς συναιδήσις, joint knowledge, consciousness, Dem.⁹

Due to its structure συνοιδα is accompanied initially with a noun or pronoun that indicates with whom the subject shares the knowledge and this can be the subject itself. So, there are reflexive and non-reflexive constructions with συνοιδα. Besides, συνοιδα constructions indicate what the object of knowing is. In other words, it indicates the kind of knowledge that is shared with others or with oneself. The shared knowledge may be neutral, but from the beginning συνοιδα constructions acquired the specific association of being a potential witness for or against the person with whom the knowledge is shared. In Isocrates’ Speeches constructions appear 24 times, and they bear almost the same meaning.¹⁰ In the following quotations, it is obvious that the slave boy knows with the suitor that there are some money transactions that are the ground for the current lawsuit. Since he is a potential witness, he is spirited away by his master:

πυθόμενος δὲ ταῦτα Πασίων καὶ εἰδός ὅτι φανερός ἦδη πράξει περὶ τῶν ἐμαυτοῦ, ἀφανίζει Κίττον τὸν παῖδα, ὡς συνήδει περὶ τῶν χρημάτων. (Isocrates, Speeches, 17 Trapeziticus, 11)

---


¹⁰ See http://www.perseus.tufts.edu/hopper/wordfreq?lang=greek&lookup=su%2Fnoida for a complete list of συνοιδα constructions in the ancient Greek literature. I cite some of the texts that indicate the meaning of συνοιδα constructions and nuances in the meaning.
When Pasion learned this and understood that I would now bring action openly about my property, he spirited away his slave Cittus, who had knowledge of our financial transactions.

When, as a result of these meetings, men of the jury all declared that Pasion was guilty of wrongdoing and of scandalous conduct, since, in the first place, it was Pasion himself who had spirited away the slave who, so I had asserted, had knowledge of the money-dealings, although he accused us of having concealed him…

In the next place, men of the jury, I can give you the reasons why he agreed to repay me the gold; for when we had been cleared of the false accusations lodged with Satyrus, and Pasion had been unable to spirit away Cittus, who had knowledge of my deposit, he understood that.

Isaeus, a contemporary of Cittus, used Συνοίδα constructions in the same way and emphasizes that those who know the facts are witnesses:

…even if we win our case, we shall always bear the stigma of having had our rights disputed, thanks to this accursed Orestes, who, taken in adultery and having suffered the treatment which befits such evil-doers, has not even so abandoned the practice, as those who know the facts can testify.

Unless stated otherwise, all the Greek texts and their translations are retrieved from Perseus Digital Library.
To prove that he remained throughout his life at variance with Cleon, I will produce as witnesses before you those who know the facts. “Witnesses”.

In addition to these types of cases, there may be cases where one can share knowledge with someone in the sense of being an accomplice in a particular deed as in the case of Sophocles’ Antigone:

 helymen δ’ έτοίμοι καὶ μόδρους αὖρειν χεροῖν καὶ πῦρ διέρπειν καὶ θεοὺς ὀρκωμοτείν, τὸ μήτε δράσαι μήτε τὸ ξυνειδέναι τὸ πράγμα βουλεύσαντι μηδ’ εἰργασμένο. (Sophocles, Antigone, 264)

We were ready to take red-hot iron in our hands, to walk through fire and to swear oaths by the gods that we had neither done the deed, nor shared knowledge of the planning or the doing.

There are similar occurrences in the Greek literature, remarkably often in contexts mentioning or suggesting compliance in conspiracies.12

A more common and specific use of συνοιδα is with a reflexive pronoun; it is in the form of συνοιδα ἐμαυτῷ and means “I know with myself” or “I am conscious of... in myself”. The content of shared knowledge refers almost always to something negative about the subject. Pierce argues that there are neutral or positive uses of συνοιδα ἐμαυτῷ. However when we examine the instances closely, we will see that there is something negative about the matter that is known by the subject itself. Pierce’s only attested positive example is from Xenophon’ Cyropaedia:

άλλ’ ἡμεῖς, ὦ ἄνδρες, μὴ πάθωμεν ταῦτα, άλλ’ ἐπείπερ σύνισμεν ἡμῖν αὐτοῖς ἀπὸ παῖδων ἀρξάμενοι ἄσκηται ὄντες τῶν καλῶν κάγαθῶν ἔργων, ἵωμεν ἐπὶ τούς πολεμίους (Xenophon, Cyropaedia 1.5.11)

But, fellow-soldiers, let us not make this mistake; but, conscious that from our boyhood on we have practised what is good and honourable, let us go against the enemy.

---

12 Thucydides, The Peloponnesian War, 8.9.3; 8.69.2, Xenophon, Hellenica, 3.3.6, Dionysius of Halicarnassus, Antiquitates Romanae, 3.30.7; 4.38.2
Pierce focuses on “ἡμῖν αὐτοῖς ἀπὸ παίδων ἀρξάμενοι ἀσκηταὶ ὁντες τὸν καλὸν κάγαθδὸν ἔργων” and claims that content of knowledge is noble and good works (Pierce 1955, 23). However, if we read the sentence from the beginning, we can see that Xenophon talks about their being conscious of a “mistake”. This mistake is explained in the former paragraph, which is indulgence in present pleasures. It may not be made by the subjects themselves but they are conscious of their tendency to indulge in present pleasures. So, the content of συνοιδα ἐμαυτῷ is a dubious conduct.

All of Pierce’s neutral examples are from Plato and Aristotle, and he claims that they use the phrase in a technical or semi-technical way and outside of an ethical context. However, in Apology 21b, Socrates is conscious of his being not wise, and in Phaedrus 235c, his own ignorance. In Symposium 216a, Alcibiades is conscious of his helplessness to resist Socrates when listening to Socrates. Aristotle also uses συνοιδα ἐμαυτῷ in Nicomachean Ethics 1095a25 to refer to subjects being conscious of their ignorance, whereas in History of Animals 618a26, he describes a condition of a cuckoo as being conscious of its inability to defend its young, i.e., of helplessness. Because of this differentiation, Pierce argues that there are ethical and neutral uses of συνοιδα ἐμαυτῷ. However, given the examination above it can be claimed that συνοιδα ἐμαυτῷ appears always in a situation when the subject reflects and evaluates her situation and then she becomes aware of something bad about herself.

There are some consequences of this type of knowledge. Xenophon states them very explicitly in Anabasis:

ὦστις δὲ τούτων σύνοιδεν αὐτῷ παρημεληκός, τούτων ἐγὼ οὐποτ` ἄν εὐδαιμονίσαμι
…and the man who is conscious that he has disregarded such oaths, I for my part should never account happy (Xenophon, Anabasis 2.5.7)

Thus someone who is conscious of neglecting oaths can never be counted happy.

According to Demosthenes if one is conscious of his own crimes in his life, then he has a life of a hare in fear and then he in trembling:

δι᾽ ἡν εὐτυχούσης μὲν τῆς πατρίδος λαγὼ βίον ἔξης δεδιώς καὶ τρέμων καὶ ἂει πληγήσεσθαι προσδοκόν ἐφ᾽ οἷς σαυτὸς συνήδεις ἀδίκοδυν (Demosthenes, Speeches, 18 On the Crown, 263)

…you lived the life of a hare, in fear and trembling and constant expectation of a sound thrashing for the crimes that burdened your conscience…

For Aristophanes, it is not easy to live with a guilty conscience:

πῶς σὺν ἐμαυτῷ τοῦτ’ ἐγὼ ξυνείσομαι, φεύγοντ’ ἀπολύσας ἀνδρα; τί ποτε πείσομαι; ἄλλ’ ὁ πολιτίμητος θεοὶ ξύγγνωτέ μοι: ἄκων γὰρ αὐτ’ ἔδρασα κοῦ τούμοι τρόπου. (Aristophanes, Wasps, 999)

And so I have charged my conscience with the acquittal of an accused being! What will become of me? Sacred gods! Forgive me. I did it despite myself; it is not in my character.

Isocrates claims that it is not possible to escape from shameful things one has done because one always knows them:

μηδέποτε μηδὲν αἰσχρὸν ποιήσας ἔλπιζε λήσειν: καὶ γὰρ ἂν τοὺς ἄλλους λάθης, σεαυτῷ συνειδήσεις. (Isocrates, To Demonicus, 1.16)

Never hope to conceal any shameful thing which you have done; for even if you do conceal it from others, your own heart will know.

It is obvious that this type of knowing is disturbing and the way to avoid it is pointed out by Xenophophon, in Apology he describes the conditions of both experiences:

ψὶ δὲ τέλος ἔχειν ἢ δίκη, εἰπὲν αὐτὸν: ἄλλ’, ὦ ἄνδρες, τοὺς μὲν διδάσκοντας τοὺς μάρτυρας ψὶ χρῆ ἐπιορκοῦντας καταψευδομαρτυρεῖν ἐμοὶ καὶ τοὺς πειθομένους τούτοις ἀνάγκη ἕστι πολλὴν ἕαυτος συνειδέναι ἢσέβειαν καὶ ἀδικίαν: ἐμοὶ δὲ τί προσήκει νῦν μεὶον φρονείν ἢ πρὶν
κατακριθῆναι, μηδὲν ἐλεγχθὲν τι ὡς πεποίηκα τι ὄν ἐγράψαντο με; οὐδὲ γὰρ ἔγονε ἀντὶ Διός καὶ Ἡρας καὶ τὸν σὺν τούτοις θεῶν οὔτε θών τισὶ καίνοις δαίμοσιν οὔτε ὄμνος οὔτε νομίζον ἄλλους θεοὺς ἀναπέφηνα. (Xenophon, Apology, 24)

When the trial was over, Socrates (according to Hermogenes) remarked: “Well, gentlemen, those who instructed the witnesses that they must bear false witness against me, perjuring themselves to do so, and those who were won over to do this must feel in their hearts a guilty consciousness of great impiety and iniquity; but as for me, why should my spirit be any less exalted now than before my condemnation, since I have not been proved guilty of having done any of the acts mentioned in the indictment? For it has not been shown that I have sacrificed to new deities in the stead of Zeus and Hera and the gods of their company or that I have invoked ill oaths or mentioned other gods.

If one knows that he bears false witness against someone innocent, then he has to face his impiety and iniquity. Whereas, as in the case of Socrates, if one knows that he has not done anything wrong then he can feel confident. In On the murder of Herodes, Antiphon makes similar observations:

εὖ δ’ ἵστε ὅτι οὐκ ἂν ποτ’ ἦλθον εἰς τὴν πόλιν, εἰ τι ξυνήδη ἐμαυτῷ τοιούτων: νῦν δὲ πιστεύων τὸ δίκαιο, οὐ πλέονος οὐδὲν ἔστιν ἄξιον ἀνδρὶ συναγωνίζεσθαι, μηδὲν αὐτῷ συνειδότα ἀνόσιον εἰργασμένω μηδὲ εἰς τοὺς θεοὺς ἡσυχήκοτι: ἐν γὰρ τὸ τοιοῦτον ἡμὴ καὶ τὸ σῶμα ἀπειρηκός ἢ ψυχὴ συνεξέσωσιν, ἐθέλουσα ταλαιπωρεῖν διὰ τὸ μὴ ξυνειδέναι ἑαυτῇ. τῷ δὲ ξυνειδότα τοῦτο αὐτῷ πρότων πολέμιον ἔστιν: ἐτί γὰρ καὶ τοῦ σώματος ἰσχύόντος ἢ ψυχὴ προμομαχεῖται, ἣγομένη τὴν τιμωρίαν οἱ ἥκειν ταῦτην τὸν ἀσεβημάτων ἐγὼ δ’ ἐμαυτῷ τοιοῦτον οὔδὲν ξυνειδὸς ἥκω εἰς ἱμᾶς.

(Antiphon, On the murder of Herodes 5.93)

Rest assured that I should never have come to Athens, had such a crime been on my conscience. I am here, as it is, because I have faith in justice, the most precious ally of the man who has no deed of sin upon his conscience and who has committed no transgression against the gods. Often at such an hour as this, when the body has given up the struggle, its salvation is the spirit, which is ready to fight on in the conscience that it is innocent. On the other hand, he whose conscience is guilty has no worse enemy than that conscience; for his spirit fails him which his body is still unwearied, because it feels that what is approaching him is the punishment of his iniquities. But it is with no such guilty conscience that I come before you.
Only those who do not have any sin can be peaceful, others who committed crimes have the worst enemy, i.e., their “conscience”.

Although there are quite a number of instances of συνοιδα constructions in the Greek literature from the 5th century on, substantives substituting for the verb phrase are very rare in the Classical period. It appears more often in the first century BC. The first definite literary use of the verbal substantive, συνείδησις is by Democritus:

ένιοι θυμήσις φύσεως διάλυσιν οὐκ εἰδότες ἀνθρώποι, συνειδήσει δὲ τῆς ἐν τοίς βίοι κακοπραγμοσύνης, τὸν τῆς βιοτής χρόνον ἐν ταραχῆς καὶ φόβους ταλαπωρέουσι, ψεύδεα περὶ τοῦ μετά τὴν τελευτὴν μυθοπλαστέοντες χρόνου. (Democritus DK 68.B297)

Some men, not knowing the dissolution of mortal nature, but conscious (συνείδησις) of evil-doing in life, distress the time of life with disturbances and fears, fabricating false myths about the time after the end of life.

Democritus claims that although people are wrong about their belief in after life, since they know their evil-doings in life they suffer and fear. They seem to believe that there will be some consequences of these evil-doings in the after-life. Here συνείδησις refers to one’s knowing his own bad deeds.

Euripides uses σύνεσις/ συνείδησις to refer to the Orestes’ sickness from which he suffers as a result of his terrible crime, which is destroying him:

Μενέλαιος τί χρῆμα πάσχεις; τίς σ’ ἀπόλλυσιν νόσος;

‘Ορέστης ἢ σύνεσις, ὅτι σύνοιδα δεῖν’ εἰργασμένος. (Euripides, Orestes, 395)

---

13 Although translators translated some of the συνοιδα ἐμαυτῷ phrases as conscience, literally they mean “I know with myself.”

14 According to MSJ both σύνεσις and συνείδησις have the same meaning but σύνεσις is the ethical term, while σύνοιδα is admitting or recognizing it. 
Menelaus: What ails you? What is your deadly sickness?

Orestes: My conscience; I know that I am guilty of a dreadful crime.

Remarkably different than Democritus, Euripides uses συνείδησις to emphasize something stronger than just the knowledge of a dreadful crime that is known by the subject himself. In this line, συνείδησις looks like a separate entity that has the power to destroy Orestes. In the following line, he says that he has grief because of this disease, in order to clarify the nature of συνείδησις. These lines from Orestes inspired many philosophical analyses later on due mainly to two reasons. First of all, Orestes is a well-known play throughout antiquity. Availability of the text makes it the subject of many treatments. Second, presentations of a familiar experience of συνείδησις with a novel form (συνείδησις) and a powerful image (like a sickness that causes grief), as a separate entity, attract the attention of many writers. However, the earliest extended analysis dates back to the 1st century, to Plutarch. Other instances of the use of συνείδησις appear in the first century BC. What can be the reason of the disappearance of συνείδησις for three centuries and re-appearance in the first century BC? Apart from lack of evidence, there must be an explanation for this fact.

My claim is that, association of an experience (being conscious of one’s own bad deeds and following disturbing feelings) with a word (συνείδησις), which implies a separate entity is a new phenomenon that needs a different understanding of the self.

---

15 Later commentators: Jerome, Olympiodorus, John Philoponus, Philostratus
16 Philodemus, Rhet. 2.140; Dionysius of Halicarnassus, de Thuc. Jud. 8.3; Ant. 8.3.1; 8.48.5; Diodorus Siculus, 4.65
Only a certain type of understanding of the self allows one to have συνείδησις. This different understanding started to emerge during the first century BC.\(^{17}\)

Before examining this new understanding of the self and its relation to συνείδησις, we need to look at the conceptualization of the self during the Classical period that allows the experience of συνοιδα ἐμαυτῷ but not συνείδησις. Indeed, it is better to start with the Homeric -Archaic age when even συνοιδα ἐμαυτῷ did not appear. This allows us to have a clearer picture of the transformation of the understanding of the self.

Homeric literature depicts the value system and world view of his era that was shaped by mythological beliefs according to which human beings have only limited power on their lives. Bruno Snell remarks that the Homeric hero performs his greatest deeds with an extraordinary power that is bestowed on him by a god. In the same manner, his great mistakes are due to the same power (Snell 1953, Ch. 1). When Agamemnon is forced to account for his unfair treatment of Achilles, he does not take personal responsibility for his conduct but attributes his actions to the madness that was visited on him by the god and blinded him to do what he did.\(^{18}\) E. R. Dodds refers to the same text

\(^{17}\) Poets like Euripides, have a deep insight that enables them to foresee new possibilities for human beings.

\(^{18}\) Nietzsche makes the same observations regarding the Homeric Greeks in the Genealogy of Morals: "But at the same time we hear and see that even this Olympian spectator and judge is far from being irritated or thinking of them as evil because of this: "How foolish they are" he thinks in relation to the bad deeds of mortal men. And the Greeks of the strongest and bravest times conceded that much about themselves—the "foolishness," "stupidity," a little "disturbance in the head" were as far as the basis for many bad and fateful things are concerned—foolishness, not sin! Do you understand that? . . . But even this disturbance in the head was a problem, "Indeed, how is this even possible? Where could this have really come from in heads like the ones we have, we men of noble descent, happy, successful, from the best society, noble, and virtuous?" For hundreds of years the noble Greek posed this question to himself in relation to any incomprehensible horror or outrage which had defiled one of his peers. "Some god must have deluded him," he finally said, shaking his head . . . This solution is typical of the Greeks . . . In this way, the gods
and observes that the phenomenon ascribing human behavior to external powers, including where mistakes are involved, commonly occurs in Homeric texts as a result of the socio-cultural context. He argues that in this context the heroes must attribute mistakes to external powers because they cannot socially afford to acknowledge them as being their own. Thus, all irrational elements in behavior are interpreted as non-human interventions. Dodds concludes that this characteristic of the epics is a proof of the absence of an internalized ethics. Irrational impulses are removed from the self and ascribed to the external cause because public esteem represents the highest value in the Homeric world (Dodds 1973, 1-27).

Dodds is right that public esteem plays a crucial role to determine one’s worth, but it is not possible to remove something that cannot be a part of that thing at first. The Homeric self is nothing more than “a working assembly of its members” (Stocking 2007, 57). The abilities such as thinking, feeling and deciding, which are the constituents of a self, are carried on by separate entities such as thumos, phrenes (sited in the lungs), kradie, etor, ker (sited in the heart) and nous. In addition, the Homeric psyche seems to designate something like the force in human beings, which flees from the body at death, rather than the site of thinking and feeling (Snell 1953, Ch. 1). Thus, we can infer that there is no single entity that can be held responsible for the thoughts and deeds of an individual. Rather there are different parts that are subject to different forces including the extraordinary ones. The Homeric self can be just a loose unity.

______________________________

then served to justify men to a certain extent, even in bad things. They served as the origin of evil—at that time the gods took upon themselves, not punishment, but, what is nobler, the guilt” (GM II, 23).
Moreover, it is not possible for the Homeric hero to feel shame when he makes a mistake because shame comes when the hero is not strong enough to prevent a degrading action against himself or his dependents. For example it is a shame for Achilles not to be strong enough to prevent Agamemnon taking away his prize. Similarly, Patroclus’ death is an even greater shame to Achilles since he fails to protect him. Hector thinks that it will be shameful if he does not fight with Achilles because it is his duty to fight for the Trojans. Homeric heroes live according to warrior ethics “where what is valued is strength, courage, the ability to conceive and execute great deeds, and where life aimed at fame and glory and the immortality one enjoys when one’s name lives forever on men’s lips” (C. Taylor 1989, 117). It is shame for them when they fail to live in accordance with these values.

Homeric heroes appeal to extraordinary forces to make sense of their misdeeds or irrational behavior not as a result of their fear of shame but because of their mythical world view which allows them to have only limited power on their lives. They are subject to the forces of the world and hence they do not feel completely responsible for their actions. Thus, it is not surprising that Homeric literature does not have any term which refers to one’s awareness for one’s own misdeeds.

Bernard Williams, in his influential book *Shame and Necessity*, criticizes Snell’s and Dodds’ “progressivist approach” that sees the ideas of the Archaic Greeks as not just different, but inferior, or rather, as Williams puts it, “primitive ideas” which have been replaced by a more complex and refined set of conceptions that define a more mature form of ethical experience. Williams argues against this claim and defends that “there are some unacknowledged similarities between Greek conceptions and our own”, so they
should not be viewed as primitive (Williams 2003, 2). However, in his detailed analysis of the Homeric Greek culture, he confirms the findings of Dodds and Snell that are discussed above. He admits that:

the idea is that the functions of the mind, above all with regard to action, are defined in terms of categories that get their significance from ethics. This was an idea that is certainly lacking in Homer. It was left to later Greek thought to invent it. It was invented, it seems, by Plato. The tripartite division of the soul in the Republic is the earliest full expression of it… (Williams 2003, 42)

Williams may be right that the Archaic Greek culture is not “primitive”, but it is definitely different and even if Homeric people are ethical agents in the “modern” sense, they do not share the view that one has to internalize norms and try to live up to the demands of one’s own conscience. This view requires a different conceptualization of the self, as Williams claims above.

Socio-cultural changes that took place during the transition period between the Archaic and the Classical ages gave rise to the emergence of συνοιδα constructions as well. The wars with Persians and the formation of the Delian League were followed by the emergence of democratic politics and the development of legal processes. These developments changed the social reality. With the rise of Athens, it became not only the economic center of the Greek world but also the hearth of socio-cultural activities. During this period both traditional conceptions about the world and universality of moral values were challenged. People’s need for new explanations to make sense of the world and the new social order was satisfied by the sophists. They not only served those who sought political power and economic wealth by teaching necessary skills but also by making man “the measure of all things” they offered a new basis for morality according
to which a person does not measure his conduct against the expectations of others, but against personal convictions of right and wrong, good and evil. As a result, while the role of external forces was diminishing, the role of powers of human beings gained prominence and a new understanding of the self started to emerge.

1.2 Plato

Plato’s dialogues are the right place to look in order to get an idea about the conception of the self in the Classical period. They reflect not only the new understanding of their time but also shape it. In the *First Alcibiades*, Socrates tries to show to Alcibiades what man is, so that he can cultivate it. He concludes that man “is nothing other than his soul” (130c).\(^\text{19}\) Socrates remarks that what they mean by man is the “individual self” (*auton hekaston*, 130d4). According to Plato, self is nothing other than the soul that is the ruler of all which belongs to a human including his body. *Republic* provides a more detailed account of the self. Socrates asserts that the ψυχή (soul) is composed of three parts; the λογιστικός (rational), the θυμός (high-spirited) and the ἐπιθυμία (appetitive) (436a-b; 580d-581d). These three parts of the soul also correspond to the three classes of a society because soul and city are parallel. They differ only in size; one is larger and the other is smaller. This is evident since the term δικαιοσύνη (justice) applies to both of them (368d-369a). Whether in a city or an individual, justice is declared to be the state of the whole in which each part fulfills its function without attempting to interfere in the functions of others (433a). Therefore, it is “appropriate for the rational part to rule, since it is really wise and exercises foresight on behalf of the whole soul, and for the spirited

part to obey it and be its ally” (441e). Towards the end of the Book IV, Socrates concludes that a just person regulates himself in a way that he puts himself in order, harmonizes the three parts of himself like musical notes and becomes entirely a unity, self-controlled, and harmonious (σώφρονα καὶ ἴημοσμένον, 443d-e).

Hence we have a picture of the true self as a self-controlled and harmonious unity that is ruled by reason. To be ruled by reason means to be ruled by the correct understanding and being able to give an account of it (534b). And in order to have the correct understanding one has to grasp the natural order. In the natural order everything has its proper place and everything is in a relation of ruling and being ruled in accordance with their nature. There is health and justice as long as each component follows the natural order (444d-e).

Taylor claims that for Plato it is not possible to be ruled by reason and be mistaken or wrong about the order of reality. So, a perfectly rational person cannot have erroneous views about the order of things or the good such as believing in a Democritean universe of accidentally concatenating atoms, or the end of life is accumulating power or wealth. The rational person not only grasps the correct order between different goals, appetites and elements in the soul, but also the order of things in the cosmos. Because he can see that at both micro and macro levels, everything is ordered for the good (Taylor 1989, 122).

Thus the Platonic self is essentially a unity under the rule of reason that is able to grasp both the natural order and its place in this order. In the light of this vision of order, the self can regulate and rule its life for the good. Given this understanding of the
self and the ordered universe, it is possible to have a better understanding of the
phenomenon of conscience and make sense of both the absence of συνείδησις and the
recurrence of συνοιδα constructions both in Plato and in his contemporaries.

The συνοιδα constructions appear 22 times in Plato’s works. They are used in
ethical contexts to denote a kind of knowledge the subject shares with himself about his
own negative traits or misdeeds. This use is compatible with Plato’s contemporaries’ use
of συνοιδα phrases. So, it looks as if Plato does not attribute a special meaning to the
experience of sharing knowledge with oneself. However, if we pay close attention to
some of his dialogues we can observe that being aware of one’s own misdeeds,
weaknesses or contradictions play a crucial role in Plato’s ethics. In fact, this awareness
is the very condition of being a good person. Once the self is conceptualized as a self-
controlled and harmonious unity under the rule of reason, it is necessary to know what is
going on in every constituent of this unity. Thus, the reason can control them and put
them in order in accordance with a vision of natural order. If one fails to detect
incompatible desires, beliefs, attitudes in himself and put them in order one will end up
with disharmony, i.e., injustice in his soul, which is equal to being sick. Therefore, in
order to be just one has to eliminate everything that distorts harmony and keeps one away
from seeing the natural order, i.e., the good. In Apology 29d-30b, Socrates defines his
duty to obey the god by caring for truth and aiming for the best possible state of the soul
and also making other people realize this duty:

Men of Athens, I respect and love you, but I shall obey the god rather than you, and while I live
and am able to continue, I shall never give up philosophy or stop exhorting you and pointing out

the truth to any one of you whom I may meet, saying in my accustomed way: “Most excellent man, are you who are a citizen of Athens, the greatest of cities and the most famous for wisdom and power, not ashamed to care for the acquisition of wealth and for reputation and honor, when you neither care nor take thought for wisdom and truth and the perfection of your soul?” And if any of you argues the point, and says he does care, I shall not let him go at once, nor shall I go away, but I shall question and examine and cross-examine him, and if I find that he does not possess virtue, but says he does, I shall rebuke him for scorning the things that are of most importance and caring more for what is of less worth. This I shall do to whomever I meet, young and old, foreigner and citizen, but most to the citizens, inasmuch as you are more nearly related to me. For now that the god commands me to do this, and I believe that no greater good ever came to pass in the city than my service to the god. For I go about doing nothing else than urging you, young and old, not to care for your persons or your property more than for the perfection of your souls, or even so much…

Socrates describes his task as the greatest good for the city because by helping people to perfect their souls, he maintains not only the right order of their souls but also the order of the city. As we have seen above order of the soul and the city are analogous for him. This analogy is not just a rhetorical tool for Plato but it reflects the understanding of the place of the individual in ancient Greece. Individuals understand themselves in terms of social units of which they are a part. The city is the most prominent social unit for Ancient Greek people. Therefore the order within the soul of the individual is closely related to the order of the city.

Plato touches on the same issue in Crito 49a-b. In order to maintain the order of the soul one must never do any kind of wrong willingly because wrongdoing and injustice is in a very way harmful and shameful to the wrongdoer. For this reason, we have to follow the truth whether the majority agree or nor, whether we must suffer worse things or will be treated more gently.
In the *Symposium*, Alcibiades exemplifies this kind of suffering. Socrates’ speech makes him aware of his own wrongdoings and suffer in consequence. He describes his experience as follows:

My heart starts leaping in my chest, the tears come streaming down my face, even the frenzied Corybantes seem sane compared to me – and, let me tell you, I am not alone. I have heard Pericles and many other great orators, and I have admired their speeches. But nothing like this ever happened to me: they never upset me so deeply that my very own soul started protesting my life – my life! – was no better than the most miserable slave’s. And yet that is exactly how this Marsyas here at my side makes me feel all the time: he makes it seem that my life is not worth living! …Socrates is the only man in the world who has made me feel shame – ah, you didn’t think I had in me, did you? Yes, he makes me feel ashamed: I know perfectly well (συνοιδα ἐμαυτῷ) that I can’t prove he is wrong when he tells me what should I do; yet, the moment I leave his side, I go back to my old ways: I cave in to my desire to please the crowd. My whole life has become one constant effort to escape from him and keep away, but when I see him, I feel deeply ashamed, because I am doing nothing about my way of life, though I agreed with him that I should… (215e-216c).

Alcibiades thinks that his being conscious (συνοιδα ἐμαυτῷ) about his life, due to the presence of Socrates, disturbs him so badly that even the Corybantes, the legendary worshipers of Cybele, who brought their own derangement, are better than him. Given Alcibiades’ account of his experience, we can conclude that if one knows the truth, which tells him the duties he has to fulfill in order to maintain the harmony of his soul, and fails to hear this voice, he suffers deeply. Plato’s description of Alcibiades’ experience is very similar to our understanding of conscience as an inner voice that distinguishes right from wrong and blames us when we fail to do the right thing. Although Socrates makes Alcibiades realize the truth about his life, he knows with himself (συνοιδα ἐμαυτῷ) that this is the case.
Plato addresses the same issue in *Gorgias*. Socrates claims that “doing what is unjust is worse than suffering it and not paying what is due worse than paying it” (474b). Since doing injustice corrupts the soul, it is the worst thing and nothing can surpass it in pain, even suffering injustice (477c-e). Similarly, not paying what is due is keeping something unjust in the soul, and it is corrupting. “Paying what is due is getting rid of the worst thing” and it is the “treatment against corruption”. Moreover, “doing what is unjust is the second worst thing. Not paying what is due when one has done what is unjust is by its nature the first worst thing, the very worst of all” (479d). Therefore, “the happiest man is the one who does not have any badness in his soul, now that this has been shown to be the most serious kind of badness” (478d-e).

For this reason:

…a man ought to accuse himself first of all, and in the second place his relations or anyone else of his friends who may from time to time be guilty of wrong; and, instead of concealing the iniquity, to bring it to light in order that he may pay the penalty and be made healthy; and, moreover, to compel both himself and his neighbors not to cower away but to submit with closed eyes and good courage, as it were, to the cutting and burning of the surgeon, in pursuit of what is good and fair, and without reckoning in the smart: if his crimes have deserved a flogging, he must submit to the rod; if fetters, to their grip; if a fine, to its payment; if banishment, to be banished; or if death, to die; himself to be the first accuser either of himself or of his relations, and to employ his rhetoric for the purpose of so exposing their iniquities that they may be relieved of that greatest evil, injustice. (*Gorgias*, 480c-d)

Therefore, Socrates concludes that “it’s better to have my lyre or a chorus that I might lead out of tune and dissonant, and have the vast majority of men disagree with me and contradict me, than to be out of harmony with myself, to contradict myself, though I’m only one person” (482b). This conclusion is also supported by Socrates’ account of a good soul in 506c-507d. A good soul is the one which is self-controlled and organized in accordance with a pre-existing order. So, a self-controlled man avoids what is
inappropriate and does not harm the harmony of his soul. These lines also foreshadow Plato’s account of the self in *Republic* which is discussed above.

This discussion reaffirms the claim that unity and harmony of the soul plays a crucial role in Plato’s ethics. Therefore, a person has to avoid any action that can distort this harmony and if he has some misdeeds he has to repair the harm in order to regain the harmony. Given this picture, it can be claimed that if one knows that he did something wrong or he has some weakness that he has to overcome, even if nobody else knows them, one experiences the same disturbance. This disturbance is explained with a metaphor in the *Hippias Major*:

Hippias, my friend, you’re a lucky man, because you know which activities a man should practice, and you’ve practiced them too – successfully, as you say. But I’m apparently held back by my crazy luck. I wander around and I’m always getting stuck. If I make a display of how stuck I am to you wise men, I get mud-spattered by your speeches when I display it. You all say what you just said, that I am spending my time on things that are silly and small and worthless. But when I’m convinced by you and say what you say, that it’s much the most excellent thing to be able to present a speech well and finely, and get things done in court or any other gathering, I hear every insult from that man who has always been refuting me (με ἐλέγχοντος). He happens to be a close relative of mine, and he lives in the same house. So when I go home to my own place and he hears me saying those things, he asks if I’m not ashamed that I dare discuss fine activities when I’ve been so plainly refuted about the fine, and it’s clear I don’t even know at all what that is itself! “Look,” he’ll say. “How will you know whose speech – or any other action – is finely presented or not, when you are ignorant of the fine? And when you’re in a state like that, do you think it’s any better for you to live than die?” That’s what I get, as I said. Insults and blame from you, insults from him. But I suppose it is necessary to bear all that. It wouldn’t be strange if it were good for me. I actually think, Hippias, that associating with both of you has done me good. The proverb says, “What’s fine is hard” – I think I know that. (304c-e)

Hannah Arendt also observes the connection between these dialogues in terms of their conceptualization of the self and interprets this passage in the following way:

When Hippias goes home, he remains one, for, though he lives alone, he does not seek to keep himself company. He certainly does not lose consciousness; he is simply not in the habit of
actualizing it. Instead, when Socrates goes home, he is not alone, he is by himself. Clearly, with this fellow who awaits him, Socrates has to come to some kind of agreement, because they live under the same roof. Better to be at odds with the whole world than be at odds with the only one you are forced to live together with when you have left company behind. (Arendt 1978, 188)

Arendt reverses Aristotle’s definition of friendship without referring to him and claims that “the self, too, is a kind of friend” in the same way “a friend is another self” (NE 1166a30). This relationship enables one to “carry on a dialogue of thought with oneself” as long as one does not contradict oneself. She underlines the impossibility of conducting a dialogue with oneself if the soul is not in harmony but at war with itself (Arendt 1978, 189). For this reason “it is better to suffer wrong than to do wrong, because one can remain a friend of the sufferer; who would want to be the friend of and have to live together with a murderer?” (Arendt 1978, 188). Arendt identifies the self as a friend or the relative of Socrates with conscience which “is the anticipation of the fellow who awaits you if and when you come home” (Arendt 1978, 191). Coming home is a metaphor for thinking in the sense of examining ourselves in terms of our deeds, thoughts, beliefs, and values. Whenever we engage in a silent dialogue with ourselves there is a possibility of encountering a contradiction within ourselves, which can distort the harmony of our soul. The fear of having a disharmonious soul and a blaming and insulting relative at home is what conscience is. It compels us to check our deeds and words in order to live peacefully with ourselves. Hence we can conclude that fear of knowing with oneself that one is inconsistent, lies at the heart of Plato’s ethics.

It is significant that unlike his fellow citizens whose ethics is shaped by their fear of external punishment and shame, Plato’s ethics is shaped by the fear of internal turmoil.
Due to his conceptualization of the self, Plato shifts the point of focus from outside to inside and this shift created a new paradigm for ethics and remained dominant for centuries. Taylor calls this transformation “internalization” (Taylor 1989, 124). By employing *elenchus* as a method of internalization, we pay attention to and can be aware of what is going on inside us and if we find out that there are inconsistencies in our actions and thoughts we feel disturbed. So, Plato not only transforms the focus of ethics but also the meaning of having knowledge about our own misdeeds and inconsistencies (*σύνοιδα* ἐμαυτῷ). Although Plato did not give a specific name to this experience his followers do not hesitate to use *suneidesis* (συνείδησις) and tried to conceptualize it.

1.3 Aristotle

Aristotle uses *σύνοιδα* constructions only a couple of times and does not really talk about the experience. Nevertheless, Aristotle plays an important role in the medieval discussions of conscience. His analysis of practical wisdom in relation with moral virtue and *akrasia* provide a framework for the later studies. This section aims to outline this framework.

Aristotle examines practical wisdom and *akrasia* in his ethics works, especially in *Nicomachean Ethics*, where he explains his understanding of eudaimonism. He claims that the ultimate end of human life is happiness. Human beings achieve happiness by

---

21 It has to be remarked that Plato’s conceptualization of the self is influenced by the external conditions of his society. Plato’s observations about the rise and decline of the Athenian democracy make him adopt Socratic eudaimonism, according to which the ultimate end of human life is happiness and it is achieved by the cultivation of virtues. In order to prevent the extreme implementations of political regimes, such as the execution of Socrates, Plato emphasizes the cultivation of virtues, through which it is possible to maintain the orderliness of the individuals and hence the orderliness of the society. So that a moderate version of democracy can flourish.

22 See, *Rhetoric* 2.5, 1382b6; *History of Animals* 9.29, 618a26; *Nicomachean Ethics* 1.4, 1095a25
fulfilling their function well. The function of a being is the defining characteristic of that being. For human beings it is “a certain kind of life, and this to be an activity or actions of the soul implying a rational principle (NE 1098a13-14). Human beings fulfill their function by exercising their reason in action and hence happiness for humans “turns out to be the activity of soul in accordance with excellence (arête/virtue)” (NE 1088a16). Aristotle distinguishes between virtues of character and virtues of intellect. Virtues of character regulate our feelings and desires with respect to pleasures and pains. Some of the intellectual virtues are related to the practical life, whereas others are more related with theoretical things, such as mathematics. The good life is attained (among other things, such as health, wealth, luck and friends) with a combination of moral virtue to direct us to the good things, and intellectual virtue (phronesis) to pick out what the appropriate action is in the particular situation.

Aristotle emphasizes that human beings learn to be good by habituation. We learn to do noble and just things by being habituated to doing noble and just things. This is the first step toward becoming a good person. The next step is to understand why to do the noble and just things and choose them for their own sake (NE, 1105a28-33). Phronesis provides such an understanding. It is the intellectual capacity to evaluate a particular situation and determine what is to be done there. Aristotle defines phronesis thus: “a true and reasoned state of capacity to act with regard to the things that are good or bad for man” (NE, 1140b6-13). Hence, phronesis is a practical deliberative ability. It apprehends both the particular goods of the situation and the universal goods in terms of which we assess the particulars. So, it grasps the ends and deliberates upon the particulars. The ends are given by moral virtue and phronesis deliberates the things with respect to these ends.
Since it is developed parallel to the development of moral virtue through habituation, it is not possible to forget *phronesis* (*NE*, 1140b21-30). Once we understand why we should desire the good things thanks to the cultivation of *phronesis*, we desire good things “naturally”. Hence we do not need to remind ourselves to do and desire the good things.

Aristotle underlines the gradual development of moral virtue and claims that in order to have complete virtue (*ἡ κυρία ἀρετή*), natural virtue (*ἡ ἀρετή φυσική*) has to involve *phronesis* (*NE*, 1144b14-17). Human beings were born with natural dispositions to do the virtuous acts but without the development of the intellect through habituation of character, the complete virtue does not develop. In the process of development of complete moral virtue, shame plays a significant role. Aristotle explains it as follows:

Shame should not be described as a virtue; for it is more like a feeling than a state of character. It is defined, at any rate, as a kind of fear of dishonor...

The feeling is not becoming to every age, but only to youth. For we think young people should be prone to the feeling of shame because they live by feeling and therefore commit many errors, but are restrained by shame; and we praise young people who are prone to this feeling, but an older person no one would praise for being prone to the sense of disgrace, since we think he should not do anything that need cause this sense. For the sense of disgrace is not even characteristic of a good man, since it is consequent on bad actions…(*NE*, 1128b10-28).

M. F. Burnyeat reads shame as a “semi virtue of the learner,” who is on his way to develop his moral virtue. Burnyeat argues that a student of moral education is already a well-brought-up person, who has been taught to enjoy noble and just actions thanks to being habituated to them. So, when he realizes that his actions are unjust or ignoble, “he feels badly about it, ashamed of his failure. The actions pain him internally, not

---

23 In the next chapter, we will revisit the idea that “we cannot forget *phronesis* as *synderesis* cannot be extinguished in the heart of men.
consequently. He is therefore, receptive to the kind of moral education, which will set his judgments straight and develop the intellectual capacities (practical wisdom), which will enable him to avoid such errors” (Burnyeat 1980, 79).

Only someone with a sense of shame will develop his virtues because he wants to do better at the right sorts of things. Hence a child would follow the directions of his tutor and learns to live according to reason and harmonize his appetites with reason (NE, 1119b13-18; 1102b28; 1166a13-14). Once he develops the moral virtues and phronesis, by which he deliberates upon a particular action by himself, then he does not need to have shame because he would not choose bad actions. Phronesis would enable him to choose the noble and just actions for their own sake. Therefore, Aristotle explicitly states that “it is clear, then, from what has been said, that it is not possible to be good in the strict sense without practical wisdom, nor practically wise without moral excellence” (NE, 1144b30-32). For this reason, Aristotle claims that Socrates is wrong in thinking that all the virtues are forms of phronesis (NE, 1144b17-25). Aristotle criticizes Socrates for this intellectualist approach, viewing virtue as the same as logos. For Aristotle virtue is a state with the right reason. It is not the same as with logos because the ends of human action are given by the moral virtue. It is not enough to have moral virtue according to reason because it might be someone else’s reason (this is the premature stage in the development of virtue, when the child follows his tutor’s reason). Only the person, who develops his own phronesis can also develop moral virtue because mindless repetition of virtuous acts does not amount to virtue. Virtue requires development of the character through deliberate choices, which is enabled by phronesis. Therefore, moral virtue interlocked with phronesis is the complete virtue.
Given the analysis of *phronesis*, it is clear that Aristotle does not need a warning or accusation mechanism for wrong-doings, i.e., conscience. In the early stages of moral development, shame fulfills this duty and in the case of a mature good man there is no need for such a mechanism because he does not do wrong voluntarily.

In the light of these explanations, Aristotle depicts a good man and a bad man in the following way:

But each of them is also found in a *good man’s relation to himself* (and in those of all other men as well, in so far as they believe themselves to be good; but, as has been said, virtue and the virtuous man seem to be the standard in everything). *For the good man is of one mind with himself, and desires the same things with every part of his nature.* Also he wishes his own good, real as well as apparent, and seeks it by action for it is a mark of a good man to exert himself actively for the good; and he does so for his own sake (for he does it on account of the intellectual part of himself, and this appears to be a man’s real self). Also he desires his own life and security, and especially that of his rational part. For existence is good for the virtuous man; and everyone wishes his own good: no one would choose to possess every good in the world on condition of becoming somebody else (for God possesses the good even as it is) *but only while remaining himself, whatever he may be; and it would appear that the thinking part is the real self, or is so more than anything else.* And the good man desires his own company; for he enjoys being by himself, since he has agreeable memories of the past and good hopes for the future, which are pleasant too; also his mind is stored with subjects for contemplation. And he is keenly conscious of his own joys and sorrows; for the same things give him pleasure or pain at all times, and not different things at different times, since he is not apt to change his mind.

It is therefore because the good man has these various feelings towards himself, and because he feels towards his friend in the same way as towards himself *for a friend is another self,* that friendship also is thought to consist in one or other of these feelings, and the possession of them is thought to be the test of a friend.

Whether a man can be said actually to feel friendship for himself is a question that may be dismissed for the present; though it may be held that he can do so in so far as he is a dual or composite being and because very intense friendship resembles self–love (*NE*, 1166a, emphasis added).
For such persons are at variance with themselves, desiring one thing and wishing another: this is the mark of the unrestrained, who choose what is pleasant but harmful instead of what they themselves think to be good. Others again, out of cowardice and idleness, neglect to do what they think best for their own interests. And men who have committed a number of crimes, and are hated for their wickedness, actually flee from life and make away with themselves. Also bad men constantly seek the society of others and shun their own company, because when they are by themselves they recall much that was unpleasant in the past and anticipate the same in the future, whereas with other people they can forget. Moreover they feel no affection for themselves, because they have no lovable qualities. Hence such men do not enter into their own joys and sorrows, as there is civil war in their souls…(NE, 1166b)

Similar to Plato’s account, Aristotle’s good man is at peace with himself, his thoughts and desires are in line with each other, he enjoys his own company like a good friend. The bad man on the other hand, has an annoying “relative at home” who reminds him of his misdeeds, and the bad man wants to stay away from this relative. The experience of the bad man described above is similar to the torments of a guilty conscience. Aristotle does not give a name to this experience but obviously this experience is due to the failure of proper cultivation of moral and intellectual virtues in addition to the lack of other “goods”, such as health, wealth, luck and friends. So, it can be concluded that one of the functions of conscience that is directing one to choose the good, is also found in *phronesis* and moral virtue.

This similarity provides a framework for the medieval studies on conscience. When medieval thinkers examine conscience with respect to its two aspects, namely *synderesis* and *conscientia*, they associate *phronesis* sometimes with *synderesis* and sometimes *conscientia*. In addition, Aristotle’s analysis of *akrasia* enables them to make sense of fallibility of *conscientia*. They employ this analysis to answer to the question: how it is possible to sin given the infallibility of *synderesis*?
Akrasia (incontinence/weakness of will) is analyzed in *NE* Book VII. According to Aristotle the incontinent man is the one who knows what one should do (has knowledge about what is right) and yet does not do it (*NE*, 1145b12). To examine this problem, Aristotle distinguishes two senses of knowing: 1) having knowledge and not using it and 2) having knowledge and using it (*NE*, 1146b30-35). So, it may be the case that even though a man knows that what he is about to do is wrong he does it anyway because he does not use this knowledge. Nevertheless this way of knowing seems strange and one has to explain what prevents a man using his knowledge.

Aristotle’s explanation concerns the relation between universal and particular propositions. It is possible for a man to know a universal proposition and either not to know the particular or not to use the particular. However, incontinence is different from making a mistake in the process of inference from a universal to a particular. It is more like a state that prevents us to infer from a universal to a particular or to use this particular which leads us to do a certain action.

Aristotle lists some states where one fails to recognize a particular or to use it such as sleep, madness, or drunkenness. In these cases one is not able to use his reasoning capacities properly so it is quite possible for him not to have the particular proposition that enables one to use his knowledge. Aristotle also claims that passions have similar effects on a human, which alters the bodily conditions so that they may behave like the mad. Aristotle remarks that the incontinent man is under similar influences since there is something that prevents him from using his knowledge. He further claims that “the use of language in an incontinent state means no more than its utterance by actors on the stage” (*NE*, 1147a24). So, we may conceptualize incontinence as a state that influences a man in
a way that he is not able to either to recognize a particular proposition or to use it. This failure may be due to either impetuosity that is being unable to complete the deliberation because of new stimuli or weakness that is being unable to be bound by the deliberations when one is influenced by desires. It is significant that the incontinent person lets himself be open to the influences of drunkenness, madness, or passions that keep his knowledge inactive, or distort his perception in a way that he cannot make sense of an object due to his appetites. So, he fails to put himself in the requisite state to meet the situation properly. In other words, the incontinent person fails to cultivate the virtues that would allow him to exercise the proper knowledge correctly. There is a failure of training that occurred long before the particular case arose. To avoid the problem of *akrasia* one has to cultivate the virtues properly.

As a result, according to Aristotle, cultivation of the virtues, especially the cultivation of practical wisdom together with the moral virtue is required in order to lead a good life. The primary aim of human beings is the fulfilment of their nature by realizing the correct order of ends in their life and integrating them into a unified whole in which each part has its proper place. This is achieved with the help of practical wisdom and other virtues. For this reason, any inconsistency or evil act is a result of inadequate cultivation of virtues. Human beings are responsible for the cultivation of their virtues, if they neglect them then they have to run away from themselves and seek the company of others “because bad men… when they are by themselves they recall much that was unpleasant in the past and anticipate the same in the future, whereas with other people they can forget. Moreover they feel no affection for themselves, because they have no lovable qualities” (*NE*, 1166b). Just like Plato’s self, Aristotle’s self is not content whenever he witnesses a
misdeed or an inconsistency of his own; however, the remedy for the Aristotelian self is to maintain the unity through the cultivation of the virtues. Therefore, he does not need to postulate a virtue or a disposition to warn and/or accuse the agent for one’s wrongdoings because practical wisdom and moral virtue develop together and there is no internal source of knowledge other than *phronesis* concerning the particulars. Nevertheless, medieval thinkers attribute some of the functions of *phronesis* to conscience, as will be shown in the next chapter.

1.4 Hellenistic Period

The golden age of Athenian philosophy continued for about a hundred years. In the following centuries, changes in the political and cultural climate of the ancient world gave rise to many varieties of philosophical thinking. During the Hellenistic period political power was vested in a highly centralized state, established and maintained primarily through extensive applications of military force. The Athenian tradition of participatory government disappeared as individual citizens were excluded from significantly shaping the social structure of their lives. Diminishing power of individuals on political affairs led people to focus more on their personal lives.

Therefore, Hellenistic philosophers, unlike their predecessors Plato and Aristotle, did not deal with the problem of the construction of an ideal state that would facilitate the achievement of a happy life. Instead, they focused upon the life of the individual by maintaining the eudemonistic framework in new forms. They tried to describe the kinds

---

24 Hellenistic period covers the period between the death of Alexander the Great in 323 BC and the emergence of the Roman Empire as signified by the Battle of Actium in 31 BC. Although some of the thinkers that are covered in this section did not live in this period, still they are considered to be Hellenistic because they share the characteristic traits of the period for the most part.
of character and action that might enable one to live well, despite the prevailing political realities. Hellenistic philosophers claim that when circumstances are beyond our control we should attempt to manage only what is for us personally manageable, our own character and thoughts. Also, we should become self-sufficient and not rely on anyone or anything outside ourselves for our well-being. Hence, we must adopt an attitude toward external events which will result in attaining peace of mind/imperturbability (ἀταραξία).

Given this framework, it would not be wrong to assert that Hellenistic thinking contributed a lot to the development of the conceptualization of conscience. Philosophers of this era tried to make sense of the new world by focusing on the inner life because little else remained but to withdraw into the self, trying to secure the inner life from external, uncontrollable influences and making happiness dependent on the state of soul alone. Although there were various Hellenistic schools of thought, Stoicism played a significant role in the development of the concept. Some scholars even argue that the term is Stoic in origin (Marietta 1970, 176-187). However, as it is argued above, the term conscience emerged in 5th century BC and continued to acquire its meaning throughout time.

The shift in focus in philosophy, during the Hellenistic period, influenced the understanding of the self, as well. Nevertheless, it can be claimed that Stoicism was a kind of continuation of classical philosophy. Plato’s portrayal of Socrates served as an exemplar of ‘self-mastery’ (ἐγκράτεια). Both this exemplar and the new sociopolitical context enabled Stoics to develop a new understanding of the ideal human being as a “self that is completely transparent to reflection, and over which its owner claims such complete authority that he finds himself in total charge of where his life is going and
indulges his emotions and appetites only to the extent that he himself determines” (A. Long 2006, 143). New philosophical ideas were continuations of previous topics but they were re-conceptualized with new terms such as \( \dot{\alpha}t\alpha\rho\alpha\varepsilon\iota\alpha \) (imperturbability),\(^{25}\) \( \dot{\alpha}p\acute{a}\theta\epsilon\iota\alpha \) (freedom from emotion), in addition to \( \omega\tau\acute{a}r\acute{r}\epsilon\iota\alpha \) (self-sufficiency), \( \dot{\alpha}r\acute{e}t\iota \) (virtue) and \( \lambda\acute{o}g\omicron\varsigma \) (logos).

Taylor remarks that like Plato, rationality is a vision of order, for the Stoics (Taylor 196, 126). The Stoics see not only the false opinions about the good but also the goodness of the whole order of things and love this order. Actually, with the help of the second one the first one is possible. This love also liberates him from caring for the particular advantageous or unfavorable outcomes which hold most men in hope and fear, pain and pleasure. For this reason, a vision of the cosmic order becomes an essential condition of true virtue (\( \dot{\alpha}r\acute{e}t\iota \)) which is representing the laws of nature.

Hence in order to lead a good life one has to free himself from all external influences and reach self-sufficiency (\( \alpha\upsilon\tau\acute{a}r\acute{k}\epsilon\iota\alpha \)). In addition, one has to seek imperturbability (\( \dot{\alpha}t\alpha\rho\alpha\varepsilon\iota\alpha \)) by living according to the demands of virtue and reason. Otherwise, behavior does not correspond with order. Such behavior does not come from the rational part of the soul (\( \nu\omicron\nu\varsigma \)), but from lower parts, the emotions and passions. Thus, the rational part has to control the soul in order to maintain the harmony.

Any transgression of the ethical code threatens the harmony of the soul by causing perturbations (\( \tau\alpha\rho\alpha\chi\iota \)- the opposite of imperturbability-\( \dot{\alpha}t\alpha\rho\alpha\varepsilon\iota\alpha \)), other signs of lack of

\(^{25}\) \( \dot{\alpha}t\alpha\rho\alpha\varepsilon\iota\alpha \) is a term that is used in the classical period but it is more emphasized in the Hellenistic period.
self-mastery (ἐγκράτεια) and undesirable emotions. Since συνειδός-συνειδήσις\textsuperscript{26} refers to an awareness of a transgression of the ethical code, it was conceptualized as one of the components of the soul.\textsuperscript{27} Although Hellenistic writers understood συνειδός as one of the components of the soul, since it is an ambiguous entity, they tried to explain it by various metaphors and similes, most of them related to the law court, e.g. judge, witness, accuser, reprover and punisher. Other metaphors were child’s nurse and disease, as they are shown in the table below:

<table>
<thead>
<tr>
<th>Writer</th>
<th>Locus</th>
<th>Metaphor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polybius</td>
<td>xviii. 43.13</td>
<td>fearful witness and terrible accuser</td>
</tr>
<tr>
<td>(c. 200 – c. 118 BC)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cicero (c. 107 – c. 43)</td>
<td>Pro Cluentio LVII, 160</td>
<td>best counsellor (optimorum consiliorum)</td>
</tr>
<tr>
<td>Diodorus Siculus (c. 90 – c. 30 BC)</td>
<td>iv.65.7</td>
<td>the punisher</td>
</tr>
<tr>
<td>Philo</td>
<td>Flaccum 7</td>
<td>the law court (dikastērion)</td>
</tr>
<tr>
<td>(c. 20 BCE – c. 50 CE)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>De Opificio Mundi 128</td>
<td>a judge (dikastēs)</td>
</tr>
</tbody>
</table>

\textsuperscript{26} Hellenistic writers used both forms. Marietta (1970, 178) and Pierce (1955, 30) agree that they refer to the same thing. Marietta argues that “syneidesis and syneidos had the same meaning, but more style conscious writers such as Josephus, Philo, Simplicius, and Plutarch prefer syneidos, while Wisdom of Solomon, Dionysius of Halicarnassus, and Diodorus Siculus use syneidesis.”

\textsuperscript{27} Dionysius of Halicarnassus, De Thucydide viii.3, Philo; De Opificio Mundi 128, Quod Deterius Potiori Insidiari Soleat 23, Josephus, Antiquities xvi.4.2
<table>
<thead>
<tr>
<th>Source</th>
<th>Reference</th>
<th>Role/Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>In Quod Deus Immutabilis</td>
<td>Sit 128</td>
<td>terrible accuser, a judge</td>
</tr>
<tr>
<td>De Decalogo 87</td>
<td></td>
<td>accuser (katēgoros)</td>
</tr>
<tr>
<td>Quod Deterious Potiori Insidiari Soleat 23</td>
<td></td>
<td>Witness (martys), accuser (katēgoros)</td>
</tr>
<tr>
<td>De Ebrietate 125, De Confusione Linguarum 121, De Specialibus Legibus iii.54</td>
<td></td>
<td>scrutinizer, reprover or convicter</td>
</tr>
<tr>
<td>Josephus (37 – c. 100)</td>
<td>Contra Apion ii.218</td>
<td>Witness</td>
</tr>
<tr>
<td>Epictetus (c. 55 – 135)</td>
<td>Schweighauser Fr. 97(disputed)</td>
<td>child’s nurse (paidagogos)</td>
</tr>
<tr>
<td>Seneca (c. 4 – 65)</td>
<td>Letters 83.1-2; 23.1-2</td>
<td>Watcher</td>
</tr>
<tr>
<td>Plutarch (c. 46 – 120)</td>
<td>De Tranquillitate Animi 476f-477a</td>
<td>ulcer in the flesh, pain, suffering</td>
</tr>
</tbody>
</table>

Since Philo (40 times) and Plutarch (71) used συνειδός and σύνοιδα constructions in their work frequently, it will be helpful to pay more attention to some of their texts in order to understand the Hellenistic conceptualization of the term.

In *Flaccum 7*, Philo writes:

I praise Flaccus, not because it is right to praise an enemy, but in order to make his wickedness more conspicuous; for pardon is given to a man who does wrong from ignorance of what is right;
but he who does wrong knowingly has no excuse, being already condemned by the court of his own conscience (συνειδός). 28

It is significant that although classical orators introduced the συνοιδα word group to denote being a potential witness for or against the person with whom the knowledge is shared, 29 it turned into the inner court of law. An explicitly external matter became internal. As it is described above, this was due to the changes in the understanding of the self during this period. Whenever humans become aware of their own misdeeds they judge themselves before the others because maintaining the unity of their inner selves is more important than getting external punishment. For this reason, Flaccus is convicted by his conscience before others. In 159-180 Philo depicts the outcomes of this conviction. Flaccus’ awareness of his guilt turned his daytime into a night (159). He was like:

(160) a very ground like men heavily oppressed, being weighed down by his calamities as if the heaviest of burdens was placed upon his neck… (167) being torn as to his soul with the memorials of his misfortunes… and being devoured with anguish, he went back home in the darkness of the night, praying, by reason of his immoderate and never-ending misery, that the evening would become morning, dreading the darkness which surrounded him was opposed to everything light or cheerful. (176) And he was continually giving way to dread and to apprehension, and shaking with fear in every limb and every portion of his body, and his whole soul was trembling with terror and quivering with palpitation and agitation, as if nothing in the world could possibly be a comfort to the man now that he was deprived of all favourable hopes… (179) [Flaccus says:] How long shall I, hardhearted that I am, bear up against such terrible calamities? I well know that I am afraid of death, since out of cruelty the Deity will not punish me violently, to cut short my miserable life, in order to load me to excess with irremediable miseries, which he treasures up against me, to do a pleasure to those whom I treacherously put to death.

29 Pp.2-4 above.
This vivid depiction of inner torments resulting from misdeeds indicates the power of συνειδός which not only makes man aware of his misdeeds but also punishes him because of the severity of this awareness. Philo describes συνειδός in *Quod Deterious Potiori Insidiari Soleat* 23, while elaborating on the “true human”:

This man, dwelling in the soul of each individual, is found at one time to be a ruler and monarch, and at another time to be a judge and umpire of the contest which take place in life. At times also he takes the place of a witness and accuser, and without being seen he reproves us from within, not suffering us to open our mouths, but taking up, and restraining, and birdling, with the reins of conscience (συνειδός) the self-satisfied and restive course of the tongue.

Following Plato and the Stoics, Philo also assumes that “the true human” is the higher part of the soul and it has to rule the lower parts. However, as a follower of a monotheistic religion, his conception of “true human” does not refer only to the rational powers of the human beings but also to the part which is aware of God’s laws. It is significant that Philo uses two pairs of words which have similar meanings to refer to the “true human”. Although the king (βασιλεύς) is at the same time a ruler (ἄρχων), he uses both of them to emphasize both the legislative (actually which reiterates the laws of God) and executive powers of a ruler. So, the “true human, dwelling in the soul of everybody”, has to make the laws and has to execute them. Similarly, he has to evaluate his own deeds both as a judge (δικαστής) and an umpire (βραβευτής), so that he can decide which one is right and which one has to win. Even sometimes he plays the role of a witness and an accuser and also a reprover. In other words, the “true human” is the ethical side of everyone that knows what is the right thing to do and how should it be done. Whenever he makes a mistake or commits a crime the “true human” testifies against himself because he knows what happened and what should have been done. Moreover, he accuses himself and leaves no room to defend himself by controlling himself thanks to the power
of συνειδός. Apparently, συνειδός is a faculty of the “true human” that is aware of one’s own misdeeds and never lets one deceive oneself. It alerts the true human by disturbing the tranquility whenever there is a transgression of a moral boundary.

Plutarch uses σύνεσις and σύνοιδα phrases in his work frequently, the quotation below is one of the most powerful depictions of how it disturbs the tranquility. He quotes from Euripides’ Orestes and elaborates on the nature of σύνεσις:

> My conscience (ἡ σύνεσις), since I know I’ve done a dreadful deed, like an ulcer in the flesh, leaves behind it in the soul regret which ever continues to wound and prick it. For the other pangs reason does away with, but regret is caused by reason itself, since the soul, together with its feeling of shame, is stung and chastised by itself. For as those who shiver with ague or burn with fevers are more distressed and pained than those who suffer the same discomforts through heat or cold from a source outside the body, so the pangs which Fortune brings, coming, as it were, from a source without, are lighter to bear; but that lament,

> None is to blame for this but me myself,

> which is chanted over one’s errors, coming as it does from within, makes the pain even heavier by reason of the disgrace one feels. (De tranquilitate animi, 19)

So, σύνεσις hurts not only like an ulcer but also does not heal easily and keeps torturing. The source of the continuing pain is the feeling of regret that is caused by reason. Since reason knows “what is the right thing to do”, once one fails to do it, σύνεσις keeps reminding that there is a wrong-doing, i.e., transgression of a moral boundary, like a wound, i.e., transgression of a physical boundary, which keeps reminding that there is an injury. Although, physical wounds and σύνεσις cause similar symptoms, since the source of the latter is internal, i.e., it is caused by the individual himself, it is much more painful. Plutarch, like Euripides adds a new aspect to the term. He implies that transgression of a moral boundary is not only shameful because of the nature of ethics but also it is
disgraceful because one fails to realize his ethical ideals. Thus, the awareness that “I am the one who made this mistake” is the cause of the sting of remorse.

Lastly, before moving to the examination of the development of conscience in the medieval period in the next chapter, we need to pay some attention to Paul’s contribution to the term. Paul uses συνείδησις (he prefers to use the verbal substantive) 14 times which corresponds to one third of the συνοίδα word group appearances in the New Testament. Both the frequency of Paul’s usage of συνείδησις and the new connotations, he adds to it, are important. Although there is a dispute about these connotations, there is no doubt that Paul’s contribution to the term is significant. Pierce claims that Paul’s contribution is fixing the contemporary Greek usage as a religious term that means “the pain suffered by man, as man, and therefore as a creature in the order of things, when, by his acts completed or initiated, he transgresses the limits of his nature” (Pierce 1955, 54). Eckstein argues that the word refers to a single entity that controls, judges and alerts to the behavior of both the self and others according to an established set of norms. P. Bosman, on the other hand, suggests that in order to make sense of Paul’s contribution, his identity as an apostle of newly emerging Christianity with Hellenistic-Jewish background has to be taken into account (Bosman 2003, 193). I agree with Bosman, especially within the context of this study, it is very important to understand the conceptual framework of συνείδησις in Paul, so that the nuances he adds to the term can be visible.

---

30 In 1 Corrintians 4,4; 8,7-12;10,25; 2 Corrintians: 1,12; 4,2;5,11; Romans: 2,15; 9,1; 13,5
31 Eckstein, quoted by Bosman in *Conscience in Philo and Paul*, p.192
Like Philo, Paul also pays attention to the inner world of man due to his Hellenistic-Jewish intellectual background that regards συνείδησις as a component of the soul. In this context, he conceptualizes συνείδησις as an inner monitor that “registers all the states of the inner person and reports them in an impartial and reliable manner” (Bosman 2003, 265). What is new in this conceptualization is actually inherent in the initial meaning of the συνοιδα phrases, i.e., sharing knowledge with another. Here, this inner entity, this component of the soul, first of all, monitors all kinds of conduct, like a witness. Second, it “registers any deviations from the norm of what is right and good, including those of the inner disposition that derives conduct” (Bosman 2003, 266). What differentiates συνείδησις from being a mere log book is this ability to distinguish right from wrong thanks to the knowledge (γνῶσις). In Corinthians 8,7-12, Paul elaborates on the relationship between συνείδησις and γνῶσις:

But the knowledge is not in everyone: some, up to now used to idols, eat (it) as sacrificial meat and their συνείδησις, being weak, gets sullied. Food does not bring us closer to the God: we lack nothing should we not eat, and we win nothing if we eat. But see to it that this authority of yours does not become a stumbling block to the weak. For if someone should see you, you have knowledge (γνῶσις), sitting to eat in an idol’s temple would not be the συνείδησις of him who is weak be “built up” to eat idol meat? For the weak person is destroyed by your knowledge, your brother for whom Christ died. By sinning against your brothers and injuring their weak συνείδησις, you sin against Christ. For this reason, if food causes my brother to stumble, I shall never eat meat, in order not to cause my brother to stumble.

In accordance with Hellenistic thinking, Paul, relates the weakness of the συνείδησις to the lack of γνῶσις. Since συνείδησις is a part of the highest part of the soul, i.e., rational part, when there is a deficiency in this part, e.g., lack of knowledge, συνείδησις becomes weak. Deficient γνῶσις weakens συνείδησις because it does not allow συνείδησις to distinguish right from wrong properly. As a result συνείδησις gets sullied. However, it is
remarkable that the pollution of συνείδησις is observed by Paul who possesses knowledge about the issue: “An idol is nothing at all in the world” and that “There is no God but one” (Corinthians 8, 4). Thus, we can identify a new aspect of Paul’s understanding of συνείδησις that not only monitors and registers one’s own conduct but also draws conclusions about another persons’ integrity, by the observation of their conduct (1 Cor. 10,29; 2 Cor. 4,2; 5,11).

Romans 2, 12-16 plays a crucial role in the development of the term, as we will witness in the next chapter, all medieval thinkers refer to the passage when they try to make sense of conscience. Paul claims that:

\[
\text{As many have sinned without the law, will perish without the law and as many as have sinned in the law, will be judged by the law. For not the hearers of the law are righteous with God, but the doers of the law will be justified. For whenever gentiles who do not have the law by nature do what the law requires, they are, without having the law, a law for themselves: these people show the requirements of the law to be written in their hearts while their συνείδησις bears witness and their reasoning in mutual debate accuse or also excuse them, on the day when God judges the secret things of the people, according to my gospel through Christ Jesus (Rom. 2, 12-16).}
\]

Reference to the law written in the heart of men is already present in Hebrew Scriptures.\[^{32}\] However, Paul’s association of inner law with conscience is a new insight. Inner law is the ultimate source of right and wrong and on the basis of this law conscience bears witness for a human beings’ deeds. Hence, the inner law becomes the content of conscience.

Paul’s conceptualization helps us to see the presumptions of previous thinkers who elaborate on the nature of conscience. First of all, according to Paul, συνείδησις

\[^{32}\text{See, Is. 51:7a; Jer. 38:33.}\]
monitors every action as a neutral observer. Only then it registers the action as right and wrong. However, it depends on knowledge of right and wrong to distinguish the actions. Although the source of knowledge is God-given inner law, if a person has deficient understanding of right and wrong then the conclusions of συνείδησις would be deficient. As a result, the person would not suffer from his wrong-doings because according to his norms of right and wrong there would be no problem. Only if a person knowingly transgresses his own moral boundaries συνείδησις alerts him and causes suffering. Thus, unlike Philo’s συνειδός, which acts like a prosecutor, a reprover or even as a judge (who knows the truth because as a rational being he can attain this knowledge - a presumption of Plato and adopted by Hellenistic philosophers) Paul’s συνείδησις is just like a processor that records everything and operates on an input file that contains the person’s ethical convictions. συνείδησις does not make independent judgements about one’s deeds but on the basis of its records it bears witness on the Judgment Day.

However, there is no room for ethical relativism in Paul’s thinking. On the contrary, he wants to emphasize the imperfect nature of human beings and the fallibility of our knowledge and also the responsibility of enhancing our knowledge that will enable us to make better judgments. He also reminds us that God, the omniscient, is the final judge and our συνείδησις will testify as a neutral recorder on the judgment day.

In addition, Paul breaks from the with individualistic ethics of Hellenistic thought according to which each person is responsible to maintain his inner harmony and reach absolute peace of mind even on the pain of detaching himself from all personal ties. By adding the ability of συνείδησις to draw conclusions about another person’s integrity, he makes everybody responsible for each other. It is our duty to increase the knowledge of
the weak ones. Definitely, Paul emphasizes the communal aspect in ethics. This is due to his Jewish background, which concerns one’s relation to others. Thus, the Pauline self knows that he is a part of community, is responsible for the other members of the community, although he has limited knowledge he has to improve his knowledge (while knowing that he cannot reach the absolute knowledge), he cannot conceal anything because his συνείδησις monitors everything and therefore, there is no room for self-deception.

1.5 Conclusion

From 5th century BC to 1st century AD, the connotations of the συνοιδα word group changed significantly due to the changes in socio-cultural context that gave rise also to the transformation of the understanding of the self. Once the look turned inward to make sense of the self, the internal experiences also became the focus of interest. In this period, the self transformed from a loose unity of different functions of the soul into a rational and responsible entity. At the same time, an awareness of one’s own misdeeds had become a component of the soul that distinguishes right from wrong and hence not only the term conscience emerged but also developed parallel to the changing understandings of the self. In this chapter, I tried to show the parallel evolution of these two concepts. Primarily, the focus is on the function of conscience: to make us aware of our own misdeeds and cause mental discomfort. However, they do not attempt to explain the mechanism how conscience works. Similarly, they do not question the content of conscience. It is assumed that given the rational nature of human beings they can have the knowledge of right and wrong. With Philo and especially Paul, this idea is replaced with a new one: God-given inner law of right and wrong. To explain the mechanism of
conscience and to explicate the actual content of conscience is the greatest challenge for the medieval thinkers. In the next chapter, I will examine their contribution to the understanding of conscience with reference to this challenge.
CHAPTER II

Medieval Understanding of Conscience

2.1 Background of the Discussion

As shown in the previous chapter, from the 5th century BC to the 1st century AD, the term conscience emerged and became a part of everyday language in the Greco-Roman world. In the medieval era, while working on the Scriptures and the writings of St. Paul, Christian thinkers appropriated Greco-Roman philosophical ideas and contributed to the development of the notion of conscience. At the same time, the analysis of the term took a new form, in terms of its two aspects namely, synderesis and conscientia. The origin of this distinction is obscure, but it played a major role in the understanding of conscience.

Origen’s interpretation of Ezekiel’s vision of a creature with four faces is the source of the medieval discussion of conscience with reference to these two concepts. The creature has a different face on each side of its head; in front a human, on the right a lion, on the left an ox, and at the back an eagle (Ezek 1:10). Origen interpreted the first three in terms of Plato’s tripartite division of the soul: the human represents the rational part, the lion the irascible, and the ox the appetitive part. Since the eagle does not have a counterpart in this division, it signifies, according to Origen, the soul’s presiding spirit that is “the spirit of man who is in him” (Origen, On Ezekiel, Homily 1.16 PG, 13.681). In his commentary on Romans, he also said that conscience is spirit.33

---

33 Origen, On Romans 2.15 PG, 14.839 “It appears necessary to discuss what the Apostle is referring to by ‘conscience,’ whether it is something substantially different from the heart or the soul. For it is said elsewhere of the conscience that it condemns and is not condemned, and that it judges man but is itself not judged. As John says, ‘If our conscience does not condemn us, we have confidence before God’ (1 John 3:21). And again Paul himself says in another passage, ‘this is our boast, the testimony of our
It is highly probable that on the basis of his reading of Origen, and those who followed him, Jerome called the fourth part, the “spark of conscience” (*scintilla conscientiae*). This is the Latin translation of a Greek term *synteresin* in his *Commentary on Ezekiel* 1.7:34

Most people interpret the man, the lion and the ox as the rational, emotional and appetitive parts of the soul, following Plato’s division, who calls them the *logikon* and *thymikon* and *epithymetikon*, locating reason in the brain, emotion in the gall-bladder and appetite in the liver. And they posit a fourth part which is above and beyond these three, and which the Greeks call *synteresin*: the spark of conscience which was not even extinguished in the breast of Cain after he was turned out of Paradise, and by which we discern that we sin, when we are overcome by pleasures or frenzy and meanwhile are misled by an imitation of reason. They reckon that this is strictly speaking, the eagle, which is not mixed up with the other three, but corrects them when they go wrong, and of which we read in Scripture as the spirit ‘which intercedes for us with ineffable groaning’ (Rom 8:26). ‘For no one knows what a man is really like, except the spirit which is in him’ (I Cor 2:11). And, writing to the Thessalonians, Paul also entreats for it to be kept sound together with soul and body (Thess 5:23). However, we also see that this conscience is cast down in some people, who have neither shame nor insight regarding their offences, and loses its place, as it is written in the book of Proverbs: ‘When the wicked man reaches the depths of sin, he does not care a damn.’ (Prov 18:3) So they deserve to be told: ‘You have acquired the face of a prostitute, you refuse to blush’ (Jer 3:3).35

---

34 Jerome’s Commentary on Ezekiel dates from the latter period of his life, when he was at Bethlehem. Jerome’s translation of the Homilies of Origen on Ezekiel is usually assigned to Jerome’s time in Constantinople, after his friend Vincentius urged him to this task in about 381. See, J.D. Kelly, *Jerome: His Life, Writings, and Controversies* (London, 1975; repr. Peabody, Mass., 1998), 76, 306.

Jerome’s text is important mainly for two reasons. First, his preservation of Origen’s understanding of conscience, as spirit influenced the subsequent thinkers. Second, from the twelfth century on, scholars tried to make sense of the word *synteresin*: the spark of conscience. They thought that there are two entities that needed to be explained in an analysis of conscience, *synderesis* (this form becomes the standard transliteration of the word because Greek ‘ντ’ is pronounced ‘nd’) and *conscientia*.

Some scholars argue that this distinction is due to a scribal error because in twenty-four manuscripts (written before the twelfth century) of the twenty-six manuscripts of Jerome, the Greek word for conscience is *suneidēsis* (Blic 1949, 146 – 147). However, some other scholars discuss the possibility that use of *synderesis* was deliberate because there was a late and (rather rare) Greek word συντήρησις and the verb τηρέω, which means to preserve, or to safeguard. The prefix συν- gives the meaning of observing or watching over oneself. It may even imply preserving oneself from wrongdoing. Not only Jerome’s word choice but also his way of describing conscience in terms of its different aspects may have given rise to the further distinction between *synderesis* and *conscientia*. On the one hand, he says that “*synteresin*: the spark of conscience, which was not even extinguished in the breast of Cain after he was turned out of Paradise, and by which we discern that we sin”, but on the other hand, he claims that “this conscience (*conscientia*) is cast down in some people, who have neither shame nor insight regarding their offences”. Even if there is a scribal error, Jerome reports that there

---

36 In Sorabji, 2014, p.60
is a spark of conscience that cannot be completely extinguished in a human being but the other functions of conscience may be mainly lost in some people. Therefore, it is legitimate to examine conscience in terms of its components, i.e., what are they, what are their specific tasks. Given the ambiguous origin, it is relatively unimportant to ask whether sympetesis and conscientia are the correct terms to identify the distinction within conscience, compared to the question whether this distinction enabled medieval thinkers to develop a comprehensive account of conscience.

This chapter aims to show that the treatises on sympetesis and conscientia can be read as the content and mechanism of conscience. So that it is possible to solve some of the puzzles, with which medieval thinkers dealt for centuries from Peter Lombard onwards.

2.2 Peter Lombard

Peter Lombard referred to Jerome’s use of “scintilla conscientiae: spark of conscience” in his discussion on “how the will (voluntas in Latin/ boulesis in Greek) could be bad”, in his Sententiae (Sentences, 1152), which became the basic textbook of his era. And consequently in the following centuries a treatise on conscience became a standard part of commentaries on Sentences.

Peter Lombard asks “since will is one of those things which man has naturally, why it is said to sin, when nothing else, which is natural, is sinful” (2.39.1, p.90). He reports that man has will naturally, just like memory and reason or thought. He lists the things, which belong to a rational creature. He claims that following Augustine, this creature is the most excellent among the created things, who is capable of changing from
a deformed to a well-formed form with the grace of God. Nevertheless, the will of this rational creature is not always good, but it is sometimes subject to defects. Some people answer that will as such is good, but when it is disordered it becomes bad and sinful.

Peter Lombard finds this answer superficial and challenges it by asking why thought, reason, and abilities are not considered to be sinful when they are disordered just like the will, by turning away from a right goal and being exercised in collusion with each other. He reports that to overcome this problem some people distinguish two senses of will, a kind of power and exercise of this power. As a power, will itself, “naturally innate to the soul, is never sinful, just like the power of remembering or of thinking; but exercise of this power, which is also called ‘will’, is sinful whenever it is disordered” (2.39.3, p. 92). This claim seems to rely on Aristotle’s distinction between potentiality and actuality. Hence, subsequent commentators discuss the possibility whether synderesis or conscientia is a potentiality or a disposition.

Once, Peter Lombard identified exercise of the will as an exercise of a potentiality, he compares it with other potentialities, such as memory, whose exercise is remembering, and thought whose exercise is thinking. Although exercise of the will may be a sin, seemingly, the exercise of other faculties is not because wanting what is bad, is bad, whereas, remembering or thinking something bad, is not. However, he comes up with counter examples, when one “remembers something bad in order to do it, and seeks to understand the truth in order to attack it” (2.39.2, p. 92). The only difference between the exercise of the will and these seems to be in their frequency but not in their kind.
Since the traditional approaches to the problem do not provide satisfactory answers, Peter Lombard formulates it in another way: “How the following is to be understood: A man, even one who is a slave to sin, naturally wants what is good” (2.39.3, p. 92). The same will seems naturally to want what is good and yet also serves sin. To make sense of this apparent conflict, Peter Lombard posits the possibility of the existence of two wills. He examines this possibility by reviewing existing claims about it. Some argue that due to a kind of mental motivation, man abandons “the law of higher things subjects itself to sins and is attracted by them”. This motivation rules over the man and suppresses the other motivation by which man desires the good, unless grace is given him. Others, however, say that man naturally wants what is good because this is the way he is created. In his original state, man has a righteous will. Since he has also free will, he may sin by his own will. For this reason, “the spark of higher reason which, as Jerome says, could not even be extinguished in Cain, always wants what is good and hates what is bad (Commentary on Ezekiel 1.7)” (2.39.3, p. 93). Although Peter Lombard did not use the term synteresis, or even scintilla conscientiae, but superior scintilla rationis, this reference became the starting point of commentaries on conscience. Since Jerome’ commentary on Ezekiel was already a part of the Glossa (the basic textbook of 12th-13th centuries), the original term he used, was available for the scholars (Greene 1991, 195–219). Moreover, in medieval theology, conscience was usually identified with reason and especially with practical reason, as we will see in this chapter. Conscience belongs to the higher part of reason, that is distinct and above the other three parts of the soul. In this sense—following Augustine, (scintilla rationis, in qua factus es ad imaginem Dei, City of God, 22.24)—Peter Lombard uses the term ‘spark of reason’ to refer to conscience to
designate its being one of the highest capacities of human reason, which is the closest image of God.

The idea that, *synderesis*, as the uncorrupted part of the soul, is incapable of error always tends towards good and away from evil, shaped the understanding of conscience throughout the late medieval era. However, there were debates about the nature of conscience. For example, whether it is a part of reason or will, and whether it is a disposition or a potentiality, etc. Starting with Philip the Chancellor, scholars tried to answer these and other related questions, in order to explain conscience. Although there is a variety among answers, basically there are two approaches: voluntaristic and intellectualistic (D’Arcy 1961, 29). The voluntaristic approach is pursued by Franciscan thinkers, especially Bonaventure, whereas the intellectualistic approach is crystallized by Thomas Aquinas. As Potts suggests, both of these traditions seem to derive from a treatise on conscience by Philip the Chancellor (Potts 1981, 31). For this reason these thinkers will be most prominently treated.

### 2.3 Philip the Chancellor

The first treatise on conscience was written by Philip the Chancellor (*Summa de bono*, 1235). His starting point is Jerome’s description of *synderesis*, “which is called the spark of conscience, which was never extinguished in Cain, and whose job is to murmur back in answer to sin and to correct mistakes” (1, p. 94). He examines conscience in four questions; in the first two questions, he determines the locus of *synderesis*, in the third, he clarifies the distinction between *synderesis* and *conscientia* with respect to their relation to sin, and finally, he inquires whether it is extinguished or not.
The first question inquires about the place of *synderesis* within the soul. Philip asks ‘is *synderesis* a potentiality of the soul or an innate disposition?’ (1.1, p. 95). Two main reasons can be identified for the formulation of the question: Aristotelian influence, via Peter Lombard’s report on the distinction between will as a power (*potentiality*) and exercise of the will (*actualization* or *disposition*), and theological doctrine, which relates *synderesis* to original sin.

Philip gives four arguments to show that *synderesis* is a potentiality. The first one is a direct quotation from Jerome’s commentary on *Ezekiel*. Since each face of the creature represents a different power of the soul; the man rational, the lion emotional, and the ox appetitive, the fourth one; the eagle, which is called *synderesis*, must be a power of the soul, as well.

The second argument also appeals to the authority of Jerome. In his commentary on *Malachi*, Jerome says “protect your spirit and the wife of your mouth” (Mal 2.15). Philip interprets this line as an evidence of natural law written in our hearts, which is wedded to our spirit. The rational spirit, as a power of the soul, has a view of and desire for God, “intercedes for us with ineffable groaning” (Rom 8.26) and it is nothing else but *synderesis*. Therefore, *synderesis* is a power of the soul.

The third argument reiterates the claim of the previous one; the spirit is a power of the soul and *synderesis* is the spirit. This time, Philip appeals to the authority of Anselm of Laon, in his commentary on I Cor 2:12. The spirit is the rational potentiality of the soul, by which we “might understand the gifts bestowed on us by God”, for this reason, we should be holy.
The first three arguments try to establish by scriptural evidence that since *synderesis* is the spirit and spirit is a rational potentiality, *synderesis* is a potentiality as well. Commentators claim that it is in our God given nature, i.e., it is a power of our natural endowment, to understand the natural law and to live accordingly. Once this potentiality is identified, it is called *synderesis*.

The fourth argument, which consists of three separate arguments, employs a different strategy and relies on “purely rational grounds” (1.4, p. 96). Each argument is based on a comparison with an already recognized potentiality of the soul. In the first one, intellect and *synderesis* are compared and Philip concludes that just like the intellect *synderesis* is also a potentiality of the soul. Intellect contemplates the truth; similarly, *synderesis* seeks good and shuns evil. On the basis of this similarity, if intellect is a potentiality then *synderesis* must be a potentiality, as well.

In the second argument, Philip puts forth the differences between sensuality and *synderesis*. Sensuality is definitely a power or the soul, which inclines reason to follow what is contingently good and to shun evil contrary to this. There will be another power, which will incline reason to follow what is good without qualification and draw back reason from what is bad without qualification. This power is *synderesis*.

Finally, Philip relies on John of Damascus’ distinction between natural and deliberative will. First he establishes that *synderesis* is not deliberative will because deliberative will does not always follow what is good naturally but sometimes acts in accordance with sin. However, natural will always desires what is good, as it is in the definition of *synderesis*, according to which it always “murmurs back in reply to sin and
corrects the mistakes”. If deliberative will is a power of the soul, then the natural will, i.e., *synderesis* will be a power of the soul.

According to Philip and his contemporaries, following Aristotle, there must be a certain power per ability in the soul. Since humans are able to think, they have a corresponding power, namely intellect or reason. Intellect has two parts; practical and speculative. The former concerns human action, whereas the latter is about eternal truths. Humans are also able to desire, for which they have appetitive powers. Similarly, appetitive power has two parts: rational and sensory. Rational appetite follows the commands of the intellect; it is called the will (*boulêsis/ voluntas*). Although thinking and desiring seem to be different kinds of abilities, unlike his contemporaries, Philip claims that they do not have separate powers in the case of practical intellect and will. On the contrary, there is a single power with two separate acts. Although the end of practical intellect is to determine the right thing to do and the end of will is good, these ends are different only intentionally, not extensionally. At the end, practical intellect determines the good and will desires it. According to Philip, since the ends of thinking and desiring have the same extension, there is only one corresponding power.

Apparently, given the argument above, according to Philip, will is a power of the soul. Since *synderesis* is a part of the will, it must be a power as well. However, in the second part of the question he tries to establish that *synderesis* is an innate disposition. Three arguments aims to show that *synderesis* is a disposition. All of them try to locate *synderesis* against a human’s tendency to sin. Since sin cannot be a potentiality of the soul, its contrary will not be either. In the first argument, the disposition to sin is associated with choice (*prohairesis*). Actually, the term *prohairesis* is Aristotelian in
origin. In *Nicomachean Ethics* 3.2-4, Aristotle describes choice as the deliberate intention with which a person acts. Choice is made about the means that contribute to the ends that are in our power, e.g. in order to be healthy we choose the acts which makes us healthy (*NE*, 1111b25-30). Medieval thinkers treat the term as a disposition of the deliberate will. Aristotle does not have a concept of will. He uses the term *boulesis* for rational desires, whose object is apparent good that could be attained by making deliberate choices (*prohairesis*) (*NE*, 1113a15-b3). Medieval thinkers adapt the same model to the relation between deliberate will and *prohairesis*: desires may be fulfilled by choosing the right means. Following Aristotelian terminology, it can be said that potentiality (*dunamis*) is the nature of a being with respect to type of change it may undergo. Actuality (*energeia* or *entelecheia*) as the realization of the change and disposition (*hexis*) is the tendency toward the way in which the change occurs in actuality. In this sense, disposition is a state in between potentiality and actuality, by which a potentiality can be actualized.

Similarly, according to medieval understanding, deliberate will is a potentiality of the human soul; it can be exercised by a disposition (*prohairesis*) that the human has acquired and directs the person to choose the means toward an end. Since both the end of the will and the means are contingently good, it is possible to make mistakes and hence to sin. Contrary to *prohairesis*, *synderesis* is always directed to what is good without qualification because it is a disposition of the natural will. As a result, since they are contraries, both *synderesis* and *prohairesis* are dispositions.

The second argument employs a similar strategy, to show that if *synderesis'* contrary is a disposition, then it will be a disposition, as well. The argument puts the impulse to sin contrary to *synderesis*. The impulse to sin inclines free choice towards sin
and evil. Given this nature it cannot be a part of human beings’ natural endowment, and a potentiality of the soul, because according to the medieval conception, potentialities are restricted to what all men have in common in virtue of being human in their original state, i.e., the state of Adam before the Fall and Christ, because they are exempt from original sin. Anything, which is an effect of original sin cannot be a substance of the soul, and hence a potentiality (Potts, p. 23 – 24). For this reason, the impulse to sin is considered to be a disposition. Synderesis is given to man by the creator to fight back against this impulse, because it is directed towards what is good. Since synderesis is an aid outside the substance of the soul (in the original state man does not need such an aid) it is a disposition, just like its contrary, the impulse to sin.

The last argument is based on an analogy between light and synderesis with respect to their role in perception, and thinking and motivating. There is a receptive potentiality for perception but without the light it is not possible to perceive things. Although light is not a part of human nature it plays a role in perception. Similarly, in thinking potentiality and motivating potentiality, there is a separate component, which inclines thought and desire toward the right direction. This component is not a potentiality, but a disposition and it is synderesis.

Once Philip reports the arguments in favor of both positions, he gives his own argument, which claims that synderesis is an innate dispositional potentiality. He agrees with all the arguments from both sides. Following Jerome and Anselm of Laon, he states that synderesis is a potentiality. Nevertheless, he argues that it is also a disposition that is contrary to prohairesis and the impulse to sin. For this reason, Philip invents the term dispositional potentiality. Thus, synderesis is given the description of a dispositional
potentiality not only out of deference to previous authors, but more importantly, to
designate it as a non-acquired, innate disposition. This compromise maintains the innate
goodness and righteousness of *synderesis*, but limits its impact, thus acknowledging the
effects of original sin. Thus, *synderesis* is, “in terms of the doctrine of original sin, what
remains after the Fall of the full control of bodily appetites which obtained before it”
(Lottin 1957, 147). In this way, *synderesis* is a little light (*modicum lumen*) leading to
God or a “murmuring back”, a whispered objection against sinful action
(*remurmurativum contra peccatum*), which thus prevents the individual from being
totally absorbed or corrupted by earthly things (Lottin 1957, 147). This is due to the
inflexible nature of *synderesis* in its appetite for the good and its hatred of evil (Potts,
1980, 28). This description in terms of an appetite or inclination to the good shows that
Philip wishes to link this innate capacity to the faculty of the will, more than reason,
although he “fudges the issue slightly by saying that it belongs to the ‘rational, not sense,
appetite’ of the will” (Hogan 2001, 68). Hogan, following D’Arcy, also suggests that
Philip’s qualification is made with the intention of linking *synderesis* to the rational,
spiritual side of human nature, rather than its animal side, with its desires and urges
(Hogan 2001, 68 - 69) since it is this aspect of the human soul that is praised in previous
Christian writing (Lottin 1957, 147). However, in this case, natural will is very similar to
reason. Since Philip is aware of this problem, in the second question, he tries to clarify
the relation between *synderesis* and reason. He asks: “is *synderesis* the same as free
choice or as reason?” (2, p. 98)

Actually, the question inquires whether *synderesis* belongs to a different category
or not, because for him free choice is a part of reason given the argument above. Hence
the question can be revised in the following way: is synderesis the same as reason? All the arguments he reports are in favor of or against this possibility. He starts with the arguments that claim it is not the same. He refers to Jerome’s commentary on Ezekiel, where synderesis is described as the fourth part of the soul. Since synderesis corrects the mistakes of the other three, including reason, it cannot be the same as reason. The source of the second argument is Augustine’s division of the soul into lower and higher parts. Both parts sin either by pleasure or consent. However, synderesis always murmurs back in answer to sin. Therefore, it cannot be the same as reason but it must be a distinct power of the soul.

Contrary to these two arguments, he lists six arguments that support the claim that synderesis is the same as reason. The first five of them appeal to the authority of the Scriptures (I Thess 5.23; Ephes 4:23; Heb 4; Matt 13:33; Job 1:15-19; Luk 10:30). The quotations can be interpreted in a way that synderesis and reason are the same because both of them are identified with the spirit, and they have the same qualities such as being an inseparable part of human nature and overseeing and correcting the other parts of the soul. The sixth argument relies on Augustine’s classification of the powers of the soul, according to which synderesis must be the power, which is reason or part of it because it has the power of commanding man to do something.

Philip’s discussion of the issue is based on his redefinition of reason. He takes four steps to locate synderesis in reason. First, he suggests that reason must be understood in a wide sense. So that reason incorporates every motivational power of the rational soul. He contrasted the rational soul with the soul in the Aristotelian sense of the life principle, which is common to all living beings. In this sense, rational soul is the soul of human
beings with all functions that are unique to their species. Since *synderesis* is one of these unique functions, it must be reason, in accordance with its wide sense.

Second, Philip tries to specify the place of *synderesis* in the rational soul, i.e., in reason. He compares *synderesis* with appetitive and emotional powers. He claims that they belong to the perceptive part, whereas *synderesis* belongs to the apprehensive part. He provides an argument to support this claim in the fourth step.

In the next step, Philip clarifies the nature of *synderesis* with respect to its origin. He claims that *synderesis* represents the reason of man in his original state, before the Fall. Then not only reason but will and emotions are directed towards good. After the Fall, the righteousness of man’s power was diminished, however, not completely taken away. What remained after the Fall is called *synderesis*. Surprisingly, Philip claims that *synderesis* is the power that exists in every other power and which “murmurs back to sin and correctly contemplates and wants what is good without qualification” (2, p.100). Jerome is right when he says that *synderesis* is beyond the three powers of the soul and above them. It is beyond them because it is inflexibly directed towards good. It is above them because of its worth, for something is worthier, which always desires good and hates evil, compared to others, which sometimes go astray. So, *synderesis* is reason, it is the inflexible part of reason and also it is part of will and emotion that are reason in its wide sense, as well.

Finally, Philip elaborates on the distinction between the higher and lower reason, a point he introduced in the second step, where he associates *synderesis* with apprehension. He claims that understanding, appetite, and emotion are motivational
powers that are to be reason. However, understanding differs from the others by lifting man above to the highest good. The other parts of reason deal with examining the features of deeds. In this case reason may be right or not right because its objects are particular things, which may be good or bad. Given this definition it is clear that understanding refers to higher reason and reason alone to the lower reason according to the terminology Philip uses. Philip’s source is Augustine, who distinguishes between the higher and lower reason. He associates higher reason with “the contemplation and consultation of things eternal” and he calls the lower reason that which “is intent on the disposal of temporal things” (*De Trin.* xii, 4). Originally, this distinction is Platonic, between knowledge and belief. The objects of knowledge are unchangeable, necessary truths. Beliefs are formed by perception and imagination of the things that are subject to change. This distinction is also akin to an Aristotelian one between theoretical and practical reason. Just like the understanding, theoretical reason is about the unchangeable, necessary truths. Practical reason, on the other hand, blends with imagination and is fed by perception is concerned with contingently good things and hence may be right or wrong. Since *synderesis* is directed always towards good without qualification, its object is unchangeable and hence it is in understanding.

Locating *synderesis* in understanding would have further consequences. As Potts remarks:

> if *synderesis* lies in understanding, then its contents must be unchangeable. Hence, the corresponding deontic propositions will be necessary and, if misunderstandings are excluded, necessarily true. General propositions, by contrast, even those which are universally quantified throughout may be contingent, e.g. ‘No one over fifty becomes an astronaut’. Universality and necessity are widely assimilated in ancient and medieval philosophy (cf. Hintikka, 1957), so Philip
probably did not realize that he was propounding two distinct constraints upon *synderesis* (Potts 1981, p. 29 – 30).

Still, Philip seems to posit that *synderesis* contains the knowledge of the good because god creates humans with this nature. Since this knowledge is provided by god, it must be unchangeable and necessarily true. This conclusion necessarily gives rise to the third question: “Can one sin by following *synderesis*?” (3, p. 102). The argument seems to be the following:

1) *Synderesis* is an innate dispositional power of reason, which directs one towards what is necessarily good.

2) Reason embraces all the powers of the rational soul.

3) So, *synderesis* is present in all the powers of the rational soul.

4) Since *synderesis* belongs to the higher reason, it is above the other powers of the soul and directs them towards what is necessarily good.

5) Therefore, human beings are able to know/reason what is necessarily good in every action by their nature (given *synderesis*).

6) If this is the case, then is it possible to sin following *synderesis*?

Philip’s solution to the problem relies on the distinction between *synderesis* and *conscientia*, whose source is Jerome’s commentary on Ezekiel as it is explained above. The argument is that one can sin by following reason because *synderesis* is the same as *conscientia*. Since, according to Jerome, *conscientia* is thrown down, ‘when the wicked man often reaches the depth of sin’, so is *synderesis*. In his reply to this argument, Philip claims that “*conscientia* comes from the conjunction of *synderesis* with free choice and is not *synderesis* itself, and is related as knowledge in action is related to knowledge in
general (\textit{universalis}) and to knowledge from reason proper, as being between them” (3, p. 104). The distinction between \textit{synderesis} and \textit{conscientia} seems to be a distinction between general and particular propositions. \textit{Synderesis} contains general necessary truths, whereas \textit{conscientia} derives particular conclusions with free choice. These conclusions may be right or mistaken. Philip’s example is as follows:

1) Everyone who makes himself out to be the son of God, and is not, should die the death. “It is written in \textit{synderesis}”.

2) This man (pointing to Christ) makes himself out to be the son of God, yet is not.

3) Therefore, he should die.

According to Philip, the conclusion is wrong because Christ is actually the son of God and hence he should not die. He states that this mistake is due to the intervention of free choice because it misapplies a general truth to a particular case and comes up with a mistaken belief. He concludes that “What was contributed by \textit{synderesis} was unchangeable and dictated only good, but this conjoined with what was contributed by reason dictated sin. So, therefore, \textit{synderesis} plus the reason for a free choice makes \textit{conscientia} right or mistaken” (3.rep., p.104). The major premise (1) is infallible \textit{knowledge} in \textit{synderesis} but the conclusion (3) is a mistaken \textit{belief} in \textit{conscientia}. And the minor premise (2) is provided by free choice because for Philip it is a matter of free choice to identify a case as an instance of a universal (in the example, identifying this man by pointing to Christ not to be the son God, is a result of free choice not an cognitive process). By introducing this distinction, Philip saves \textit{synderesis} from mistakes, i.e., sin. However, his account needs further explanations.
The nature of general propositions in synderesis is unclear. As Potts rightly asks “how general must a general deontic proposition be in order to qualify as a possible object of synderesis?” (Potts 1981, p.14). Even the simple verb has to be combined with some proper names or definite descriptions in order to form a meaningful sentence. There are some very general propositions such as “do not kill” or “do not steal” but there are some more specific general propositions as well, just as Philip’s own example above. In the first question, he associates synderesis with the natural law written in the heart but it is not clear what is written in the heart. For Philip, the proposition “Everyone who makes himself out to be the son of God, and is not, should die the death” is written in the synderesis but he does not give an account why this is necessarily so. Hence Philip’s depiction of synderesis as “knowledge in general” is not a convincing one. He fails to provide an account about the content of the general propositions in synderesis.

However it is still possible to offer an alternative interpretation for the claim that synderesis has the nature of theoretical knowledge. Although Philip says quite explicitly that the relation between synderesis and conscientia is similar to the relation between general/theoretical and practical knowledge, synderesis may not contain all the necessarily true general propositions, but it can be a cognitive ability of reason to detect what is necessarily true. However, even in this case, synderesis needs at least a basic understanding or a criterion of the good, so that it can apply it to a proposition to check whether it is true or not. Maybe for this reason Philip describes synderesis as the innate (because he assumes that there is an innate sense of the good) dispositional power (an ability to apply this sense to the general propositions). Yet, Philip has to give an account for this criterion. For example, if it is; obey the laws of God or follow the scripture, then
the next questions will be which laws or whose interpretation.\footnote{As we will see in the next chapter, Kant’s universal law can be interpreted as an attempt to provide the necessary criterion for synderesis.} Philip does not provide an explanation for these issues. Therefore, even if this interpretation (synderesis as an ability to detect or formulate general propositions) is true, it fails to give a substantial account of the good.

Suppose that the issue of providing synderesis with the criterion for the good or all the necessarily true general propositions is resolved, then for each particular circumstance, free choice has to derive some conclusions. For, synderesis is too general to be applied as it is and a decision has to be made whether a minor term is an instance of the major term in the synderesis. In other words, minor premises are determined by free choice. Therefore, conscientia, i.e., the conclusion, which follows from the major general proposition and the minor premise, can be right or mistaken. Although synderesis inclines free choice by telling it to do good and restraining it from evil, still this inclination is effective only up to a certain degree because the good in question is general and needs to be particularized. During the process of particularization, free choice is also affected by other inclinations, which may divert it from the good. Hence, synderesis is not mistaken but conscientia is.

Philip seems to resolve the initial problem in question three, which is “can one sin by following synderesis?” by introducing the distinction between synderesis and conscientia. Although he does not provide a satisfactory explanation how synderesis possesses the necessarily true general propositions, by ascribing such a nature to it, he
saves it from mistakes, i.e., sin. *Synderesis* provides reason with infallible content so that it makes judgments both retrospectively and prospectively about the rightness of a deed. Conscience works with syllogisms, where *synderesis* serves as the first premise (general proposition) and *conscientia* as the conclusion.

Finally, Philip’s first attempt at a treatise on *synderesis* results in a series of compromises, caused by different motivations, including deference to past sources. The importance of Philip the Chancellor is that his comments set much of the groundwork to future writing on *synderesis* and *conscientia*. D’Arcy considers Philip’s conclusion as “disappointing” (D’Arcy 1961, 28), whereas, Lottin considers Philip’s work to be progress in comparison to his predecessors (Lottin 1957, 148). Potts, on the other hand, claims that Philip’s treatise opened up a series of questions relating to conscience, but it did not provide definitive solutions to them (Potts 1980, 31). I believe the importance of Philip’s treatise on conscience lies in his attempt to specify the nature of the content of *synderesis*. He realizes that in order to be infallible the content of *synderesis* has to be necessarily true. However, he fails to show how general knowledge in *synderesis* is necessarily true.

2.4 Bonaventure

Twenty years after Philip’s treatise, a treatise on conscience became a customary part of commentaries on Peter Lombard’s *Sentences*. Like his contemporaries, Bonaventure (*Commentary on the Sentences*, Book II, distinction 39, circa 1230) divides his treatise into two parts; the first one is on *conscientia* and the other one is on *synderesis*. The first question in each part aims to specify the place of the term: does it belong to the thinking
or desiring part of the soul? The other questions discuss the relation to sin of *synderesis* and *conscientia*.

The answer to the first question in the *conscientia* part, regarding to which part of the soul conscientia belongs, observes the distinction between *synderesis* and *conscientia*. *Conscientia* belongs to the thinking part of the soul and *synderesis* to the desiring part. Bonaventure claims that given the different senses of *conscientia*, it can be concluded that it is a disposition of the potentiality of apprehension (*cognitio*). Since it means;

i) the thing of which we are conscious, i.e., the natural law.
ii) the potentiality of being conscious, i.e., the ability of being conscious of the natural law.
iii) the disposition to know the natural law (the more usual meaning).

Hence *conscientia* enables human beings to be conscious of the natural law and directs them to act in accordance with it. However, he remarks that the knowledge in question is not theoretical but practical because it concerns deeds. It has also a motivational aspect because it “tells us to do something and turns us towards doing it” (dist.39,a.1.q.1,dis., p. 111). Since knowledge is combined *with* desire and deed, it takes the name *con* (with) + *scientia* (knowledge), according to Bonaventure.\(^{39}\) Because of this nature *conscientia* dictates premises such as “God is to be honored” and not like “every whole is greater than its parts” (dist.39, a.1.q.1,dis.,p. 111). These premises are like the rules for what is to be done, according to Bonaventure. He concludes that *conscientia* is in the apprehensive

\(^{39}\) Definitely, *conscientia* does not imply this meaning. See chapter 1 for the meaning of the term. Probably, Bonaventure defines *conscientia* in this way for the sake of rhetoric.
part of the soul and in the capacity of practical reasoning. He refers to Aristotle to justify his claim that an apprehensive disposition could be practical because "speculative and practical thought are said to be the same potentiality, differing only in their extension" 

\textit{(De anima} 3.7.431b10-12).\

Bonaventure addresses the issue of the source of the rules in the \textit{conscientia} in the second question: Is \textit{conscientia} an innate or an acquired disposition? How do human beings know that they should honor God? Is this idea innate to the soul or is it acquired? Bonaventure appeals to the authority of Aristotle and Augustine and discusses the different aspects of the issue with reference to them. He claims that they would agree that apprehensive dispositions are like virtues. Since virtues are a matter of habit, they are neither wholly derived from nature nor wholly acquired, but partly innate and partly acquired, and so is \textit{conscientia}, as an apprehensive disposition. However, there would be some disagreements about how they are innate and acquired.\textsuperscript{40}

First, Aristotle states that apprehensive dispositions are innate with respect to the active intellect but acquired with respect to the possible intellect, and about the possible intellect he says that the soul is created like a blank sheet of paper and that this intellect is perfected by means of perceptive powers \textit{(De anima} 3.4.429b30-430a2). Bonaventure, however, finds this explanation problematic, because if the active intellect had apprehensive dispositions, it would have knowledge from the beginning, which is not the case.

\textsuperscript{40} Langston misreads the distinction made by Bonaventure, and so, although he recognizes that "conscience is divided into two general parts by Bonaventure," he declares both of them to be innate. See Langston, \textit{Conscience and Other Virtues}, 25.
Second, Bonaventure evaluates Philip’s solution. Bonaventure reports that according to Philip, apprehensive dispositions are innate with respect to apprehension in general and acquired with respect to apprehension of the particular. In other words, apprehensive dispositions are innate with respect to the apprehension of premises and acquired with respect to the apprehension of conclusions. Hence apprehensive dispositions have the nature of self-evident axioms. However, Bonaventure strongly rejects this approach with the help of Aristotle and Augustine. According to Bonaventure, Aristotle proves that apprehension of premises cannot be innate in *Posterior Analytics* (2.19.99b22ff.). It is acquired from perception, memory, and experience. Similarly, Augustine rejects the possibility of possessing innate knowledge in his interpretation of Plato’s *Meno*. The slave answers all the questions about geometry not because he recollects what he had known before birth but rather “he saw these things by a sort of incorporeal light of an unique kind; as the eye of the flesh sees things adjacent to itself in this bodily light, of which light is made to be receptive, and adapted to it” (*On the Trinity* 12.15).

Bonaventure proposes a new way of explaining how apprehensive dispositions are partly innate and partly acquired by synthesizing Aristotle’s and Augustine’s accounts:

Since it is necessary to apprehension that two things should be present concurrently, namely what can be apprehended and light by means of which we judge the former, as we see in the case of sight and as Augustine suggests in the passage quoted above, apprehensory dispositions are partly innate because of a light imparted to the soul, but also partly acquired because of forms. This accords with the words both of Aristotle and of Augustine. For everyone agrees that there is an imparted light of the apprehensory potentiality which is called a natural tribunal, but we acquire forms and likenesses of things by means of the senses, as Aristotle says explicitly in many places (*Posterior Analytics* 1.18; 2.15; *De anima* 3.8; *Metaphysics* 1.1) and as experience also teaches us. For no one would ever apprehend whole or part, or father or mother, unless he received its form
through one of the external senses; …However, that light or natural tribunal directs the soul itself in judging both of what can be apprehended and of what can be done (dist.39,a.1,q.2,resp., p. 113).

Bonaventure tries to solve the problem for which Philip does not have a satisfactory answer. How do we have the content of the general propositions or following Bonaventure’s terminology, the content of the rules for what can be done? He claims that the concepts with which we formulate the premises of theoretical or practical knowledge are acquired by means of the senses. For this reason, in terms of its content conscientia is an acquired apprehensory disposition. However, the truth value of the premises can be seen by the light of reason, which is the natural/innate tribunal to judge whether a premise is true or false. Hence, the prosecutory part of conscientia is innate.

Bonaventure remarks that not all premises are known innately. He states that some of the objects of apprehension are “exceedingly plain, e.g. axioms and primary premises” of theoretical reasoning but particular conclusions are not so plain. Similarly, some of the rules for what can be done are “maximally plain, e.g. do not do to others what you do not want to be done to you, that one ought to submit to god, and so on”. He argues that there is no big difference in the apprehension of the basic premises of theoretical and practical reasoning. The basic premises of theoretical reasoning are apprehended by virtue of the innate light, which makes these premises visible whether the claim in question is true or false, as long as the concepts in use are clearly understood. The basic premises of practical reasoning are apprehended by a similar analysis. For example, if one has a clear understanding of the concepts; god, obligation and submission, she can easily apprehend that the proposition “one ought to submit to God” is true with the natural light of reason. However, for the apprehension of the particular
conclusions of the theoretical knowledge, this natural light is not enough but some
“persuasion and new aptitude” is required. The same is true for the particular conclusions
of the practical knowledge. Bonaventure claims that they can be apprehended by
“additional education” (dist.39, a.1.q.2, resp., 113-114).

Although Bonaventure’s clarification about the extent of innate knowledge
reminds us of Philip’s distinction between general and particular propositions, actually
Bonaventure has a novel notion. For Philip general propositions themselves are innate.
However, for Bonaventure what is innate is the way we justify them. By the inner natural
light of reason, we see their truth. The light makes us see but not check their truth value
with respect to something more general. If we apply the modern terminology, it can be
said that the natural light of reason means conceptual analysis. When we have a clear
understanding of the concepts and the basic premises, which are constructed with these
concepts, they can be evaluated correctly by the natural light of reason, i.e., the innate
part of conscientia. For this reason, conscientia is infallible regarding the apprehension of
basic general premises, where no further reasoning is required. However, when things get
complicated the natural light of reason is not enough to illuminate the way from generals
to the particulars. Hence, it has to be accompanied by the deliberate reason “whose job is
to distinguish one thing from another and compare one thing with another” because, in
the case of particulars, things come together and they are not plain any more. For this
reason, mistakes may arise, as the following example indicates:

The conscientia of the Jews first told them itself by natural pronouncement that God is to be
obeyed, and they assumed henceforth that God now directs circumcision and keeping certain foods
separate. From this their conscientia is formed in the particular matter, that they should circumcise
themselves and abstain from certain foods. This mistake does not come from the first premise,
which was indeed true, but from adding the minor premise, which was not from conscientia as a natural tribunal, but rather from mistaken reason, which has regard to free choice (dist.39, a.2.q.3, resp., p. 120, emphasis added).

The truth of the premise ‘God is to be obeyed’ is indisputable, according to Bonaventure. He claims that Jews are mistaken in believing that circumcision and abstaining from certain foods are still obligatory. As a Christian he believes that there is no such an obligation anymore. Hence Jews’ beliefs are obviously wrong. However, this is a highly problematic conclusion because it implies that only “Christians” (followers of a certain interpretation) may obey God in the right way. Theological evaluation of the argument is out of the scope of this study but a philosophical analysis can be provided.

There are two competing arguments:

1. i. God is to be obeyed.
   ii. God commands all that male children have to be circumcised.
   iii. Therefore, male children have to be circumcised.

2 i. God is to be obeyed.
   ii. God commands all that male children do not have to be circumcised anymore.
   iii. Therefore, male children do not have to be circumcised anymore.

According to Bonaventure, both 1i and 2i are true because their truth can be known by the light of reason (innate part of conscientia), with an analysis of the concept of God. However, Bonaventure fails to see that even the concept of God is not immune from disputation. As a member of highly homogeneous religious society, he may share his analysis with the other members of the society. Nevertheless, this situation does not
guarantee the unanimity of the concept. Any two people may have a different understanding of the concept of god as a result of their analysis. Hence, what is obvious truth for him may not be so, for another person. The problem is that an obvious truth is not equal to a self-evident truth. A self-evident truth does not need further evidence or demonstration to be accepted as true. Its truth value does not depend on the individuals. Whereas, an obvious truth is “obviously true for someone”. People’s consent is required in order to claim that something is an obvious truth. However, Bonaventure assumes that his understanding of God is self-evident and true for everyone no matter whether they admit it or not. If they pay attention they will also see the truth of his understanding of God. Failures to have the right understanding of God are due to lack of faith, according to Bonaventure. So, the light of reason works properly when it is empowered with faith.41

It is also possible to interpret the natural light of reason as an intuition, i.e., a practical insight, which enables one to “know” the basic general premises, from which one can reason to more specific ones (Potts 1981, 38 – 41). Potts points out the central difficulty of this view “that it does not supply us with a method of identifying basic general propositions, since any two people may differ over which general propositions they claim to know by intuition” (Potts 1981, p. 38). He reviews a possible objection to this claim that is Bonaventure actually does have an independent criterion for basic general propositions, albeit a theological one. The criterion to identify the basic general

---

41 See Z. Hayes 2013, 245: “Reason is subordinate to faith; but faith is the intrinsic perfection of reason. For if the task of reason is to know the reality, and if reality is not known until it is known in Trinitarian terms, then the task of reason is realized only partially until it finds its completion in faith (III Sent. d.24, a. 2, q. 2).”
propositions is” the law of God” that amounts to natural law and simply means what is right and wrong independently of any legislation. Potts observes that:

The great majority of mankind is probably intuitionist in practice: most people think it is obvious that certain kinds of action are right and others wrong, though perhaps more perplexity is to be found today than in former times. It is therefore of considerable importance to ask how an intuitionist position, even of the medieval type proposed by Bonaventure, might be justified. One way of identifying the basic deontic propositions would be to take them as the highest common factor of everybody’s claimed intuitions. In a relatively homogeneous society, this solution has often been adopted unreflectively. Thus medieval writers are able to speak of the “law of God” or of “the natural law” as though it will be obvious to their readers what it comprises (Potts, 1981, p.38 emphasis added).

The problem still persists because the whole argument relies on the claim that the basic general propositions are obvious for certain people. As discussed above its being obvious does not guarantee the truth of a proposition. So, it can be concluded that even the truth of 1i and 2i is disputable because Bonaventure fails to provide a satisfactory account for the innate and hence infallible part of conscientia. The light of reason, whether as an ability to analyze basic general propositions or as an intuition to know them, does not guarantee the truth of a basic general proposition.

In the last question, Bonaventure examines the possible outcomes of following one’s conscientia. He states that conscientia leads one in three ways by telling one 1) what is in accordance with the law of God or 2) what is in addition to the law of God or 3) what is against the law of God. In the first case, conscientia is binding because it prescribes the basic general propositions that follow from the divine law. In the second case, conscientia derives some conclusions from the basic general propositions with the help of deliberative reason. In this case, it is binding as long as it sticks to the law of God but it is not binding and one has to change her conscientia when it deviates from the law
of God. In the third case, one is only bound to change his conscientia to make it align with the law of God. Otherwise, it would be a mistaken conscientia. Since conscientia is only infallible regarding the basic general rules, one needs a more truthful guide to follow.

According to Bonaventure, this guide is synderesis because it always desires good and directs conscientia in the same direction. It is the spark of conscience because “…conscientia, in itself, cannot move or vex or stimulate without the mediation of synderesis, which is like the stimulus and flame [of conscientia]. Thus, just as reason cannot move without the mediation of the will, so conscientia cannot move without the mediation of synderesis” (945b Translated by Langston 2001, pp. 29 – 30). Synderesis is the spark of conscience in the sense that it “provides a spark that conscientia needs to operate” (Langston 2001, p. 30). So, synderesis drives conscientia to formulate the right principles regarding the good behavior. By itself synderesis is directed toward good in general, conscientia provides the knowledge of particular good by apprehending the basic deontological propositions and deriving further consequences from them. In this process, synderesis sets conscientia in motion and shows the direction. Because of its motivational role, synderesis resides in the desiring part of the soul and it is an innate dispositional potentiality of will. According to Lottin, although Bonaventure moves synderesis from reason to the will, still he feels the necessity of bowing to the authority of his predecessors, especially Philip. So that he employs his formula when describing synderesis as a dispositional potentiality (Lottin 1957, p. 207).

Despite the guidance of synderesis, conscientia can still be mistaken. Most of the time mistaken conscientia is due to the errors in reasoning but it may be the case that
synderesis stops functioning. In the second question of the second part of the treatise, Bonaventure investigates this possibility. He claims that although the exercise of synderesis may be prevented, it cannot be extinguished because of their natural endowment, human beings are always directed to good in general. However, ‘the darkness of blindness’, or ‘wantonness of pleasure’, or ‘the hardness of obstinacy’ may prevent synderesis from functioning temporarily.

These impediments are the three possible sources of evil doing. In the first case, “synderesis does not murmur in reply to evil because the evil is believed to be good e.g., in the case of heretics who, while dying for the impiety of their error, believe that they die for their piety of faith, so that they feel no guilt, but, instead a fictitious and vain joy” (dist.39, a.2.q.2, resp., p. 117). Because of ignorance, the heretics mistakenly believe what is evil. There is a failure in their apprehension. However, since their intention is to pursue the good, they do not feel guilt.

In the second case synderesis “is hampered by the wantonness of pleasure, for sometimes in sins of flesh a man is so engrossed by the exercise of the flesh that a sense of guilt has no place, because men of the flesh are so far carried away by the impulse to pleasure that reason has then no place [in them]” (dist.39, a.2.q.2, resp., p. 117). In this case, pleasure rules over the soul, so that reason is not able to identify the object of pleasure as something evil. Due to the failure in the identification of evil, synderesis does not react to it.

In the case of the hardness of obstinacy, synderesis “does not goad towards the good, as e.g. in the case of the damned, who are so strongly reinforced in evil that they
can never turn towards what is good” (dist.39,a.2.q.2,resp., pp. 117 – 118 ). Since the obstinate repeatedly commit the same sins, their apprehension of the good changes throughout time in a way that for them evil becomes the good. So, it is not possible for them to see the good as it is anymore and their synderesis becomes ineffective. Since synderesis is inactive in all the cases, it cannot guide conscientia to change when it is mistaken.

It is remarkable that although Bonaventure assigns synderesis a motivational role, i.e., to move conscientia, actually synderesis depends on the judgments of conscientia. Conscientia decides what has to be done by formulating general rules on the basis of basic ones; however, the existence of these rules by themselves is not enough to take action on them. According to Bonaventure, one needs to want to execute them and synderesis makes us want to execute the rules, which are formulated by conscientia with the belief that they serve the good. However, synderesis has no other source but conscientia to trust about what is good in particular. As it is shown above, when heretics mistakenly believe the evil to be the good, synderesis does not say anything because it follows the judgment of conscientia. By placing conscientia (apprehensive) and synderesis (desiring) in the different part of the soul, Bonaventure limits their power. On the one hand, it looks like synderesis sets conscientia in motion and directs it toward good but synderesis depends on the judgments of conscientia about the good in particular and if conscientia fails to see the truth synderesis cannot correct it. On the other hand, conscientia decides a certain action is obligatory but without synderesis this decision is not enough to take the action. Moreover, as it is concluded above, according to Bonaventure, failures in apprehending the good are due to lack of faith. So, neither
conscientia nor synderesis works properly unless they are empowered with the “true faith”. This requirement may be the weakest point of Bonaventure’s account of conscience. Only the ones who have “true faith” can have a chance to save themselves from wrong-doing.\footnote{S. P. Chalmers (2013, p.100n139), also comes to the same conclusion: “Bonaventure does not dispute the value of reason, though he points out its limits. As a result, he declares that it is impossible not to fall into error without the light of faith. See St Bonaventure, \textit{Collationes de Septem Donis Spiritus Sancti}, in \textit{Opera Omnia}, vol. 5 (Quaracchi, Florence: St Bonaventure College, 1891), coll. IV, n. 12.”}

The most important reason to distinguish conscientia and synderesis is to save at least a part of conscience from failure. However, on the one hand, even if we accept Bonaventure’s explanation, conscientia is only infallible regarding the basic general rules. By making synderesis a part of the desiring part of the soul and claiming that it always wants us to do the good, he saves it from fallibility but it does not function when conscientia mistakenly judges something evil to be the good due to the errors in reasoning. Synderesis also is not a reliable source to save one from wrong-doing because its infallibility is too abstract. Therefore, Bonaventure’s distinction is not a successful one; neither conscientia nor synderesis is infallible.

Despite the weaknesses of his account of conscience, still Bonaventure’s contribution to the understanding of the term is important. First, he tries to solve the problem of the nature of content of conscience. What kind of knowledge should be in conscience so that one can do the right thing? Bonaventure suggests that the answer of this question has to be in the form of rules and they should be based on some necessary truths. Although he fails to provide a criterion to identify the necessary general propositions, given the inadequacy of Philip’s account on the distinction between general
and particular propositions, Bonaventure’s introduction of the necessity is a positive contribution (Potts 1981, p. 44).

Second, Langston remarks that by placing the *synderesis* in the desiring part, Bonaventure provides an answer to the question: why should conscience be followed?

Bonaventure sees the formation of ethical rules by conscience as an implementation of a human being’s desire for the good (the *synderesis*). He also sees the following of these principles as another aspect of the desire for good. Because we naturally have a desire for the good, we also desire the means to that goal. The principles of conscience are such means, and so we are naturally disposed to carry out the principles of conscience. Similarly, the emotional reaction to doing evil (guilt or remorse) is a reaction to the frustration of the desire for good caused when one fails to adhere to what the conscience has determined leads to good (Langston 2001, p. 35).

It is important that Bonaventure takes emotions into the consideration in his account of conscience, which is a promising track to pursue. However, as discussed above there are some weak points in this account. There are some gaps in the interaction between *synderesis* and *conscientia*. Moreover, it is possible to give a purely intellectualistic answer to the question as we will see in Thomas Aquinas.

2.5 Thomas Aquinas

Thomas Aquinas treats conscience in three places: in his commentary on the *Sentences* of Peter Lombard (1253-1255), in the *Disputed Questions on Truth* (*De Verit.*), 16th and 17th questions (1257-1258), and finally in the *Summa Theologiae* (*ST*) I.79 12th and 13th articles (circa 1267). In the last one he summarizes his findings in the *De Verit*. Because of the richness of the material, this study will mainly rely on the *De Verit*. In addition, since Thomas’ treatise of the natural law in *ST* supports the arguments about *synderesis* and *conscientia*, they will be taken into the account.
Following the tradition, Thomas also divides his treatise into two parts, and starts with *synderesis*. The first question in this part is about the nature of *synderesis*, as expected: is it a potentiality or a disposition? As we have seen above, Bonaventure ascribes to *conscientia* that which was assigned to *synderesis* by Philip the Chancellor. Thomas follows Philip’s differentiation more strictly and claims that *synderesis* is a dispositional potentiality of reason by which basic general propositions are known without reasoning. He gives an argument to support this claim, while Bonaventure finds it adequate to describe this disposition as the light of reason. Thomas states that given the hierarchy of beings, human beings are below the angels who are capable of apprehending the truth without inquiry:

…a lower nature, at its highest, comes near to what is proper to a higher nature, participating in the latter imperfectly…[Therefore] the human mind, at its highest, comes near to something of what is proper to an angelic nature, i.e. by apprehending some things immediately and without inquiry although, in this, it is inferior to an angel because it only apprehends the truth in such cases through the senses…Moreover, this apprehension must be the source of all subsequent apprehension, whether theoretical or practical, since sources should be more stable and certain. So this apprehension must be naturally present in man, because he apprehends it as a kind of seed-bed of all subsequent knowledge, just as natural germs of subsequent behavior and effects pre-exist in every nature. This apprehension must also be dispositional, so that it will be ready for use when needed….Accordingly, just as there is a natural disposition of the human mind by which it apprehends the principles of theoretical disciplines, which we call the understanding of principles, so too it has a natural disposition concerned with the basic principles of behavior, which are the general principles of natural law. This disposition relates *synderesis*: it exist in no other potentiality but reason… (*De verit. q.16, a.1c., p.124*).

In other words, it is in human nature to apprehend some basic theoretical and practical propositions without reasoning and *synderesis* is the name of a disposition by which human beings turn towards the good with the apprehension of the basic principles of behavior. Thomas presupposes that because of their rational nature, human beings
apprehend the first principles of natural law without inquiry and on the basis of it they can derive other principles of behavior. He does not simply say that it is written in our heart but provides an argument for how human beings apprehend them.

Since Thomas is looking for basic principles, which will serve the ground for the further reasoning, he would like to show that they are necessarily true. For this reason, he defines the criterion, which makes a principle necessarily true, i.e. being self-evident. He might be aware of the confusion between self-evident and obvious truths, which is discussed above. So, he is very cautious when he describes the term self-evident:

Now a thing is said to be self-evident in two ways: first, in itself; secondly, in relation to us. Any proposition is said to be self-evident in itself, if its predicate is contained in the notion of the subject: although, to one who knows not the definition of the subject, it happens that such a proposition is not self-evident. For instance, this proposition, "Man is a rational being," is, in its very nature, self-evident, since who says "man," says "a rational being": and yet to one who knows not what a man is, this proposition is not self-evident. Hence it is that, as Boethius says (De Hebdom.), certain axioms or propositions are universally self-evident to all; and such are those propositions whose terms are known to all, as, "Every whole is greater than its part," and, "Things equal to one and the same are equal to one another." But some propositions are self-evident only to the wise, who understand the meaning of the terms of such propositions: thus to one who understands that an angel is not a body, it is self-evident that an angel is not circumspectively in a place: but this is not evident to the unlearned, for they cannot grasp it (ST 2 -1.94.2c).

Thomas’ remark on self-evident knowledge foreshadows Kant’s definition of analytic judgments. Any proposition, which can be known true in virtue of its meaning alone, is necessarily true. So, if Thomas can show a principle of behavior in the same fashion, he will be justified in claiming that this principle is necessarily true.

---

43 Translated by Fathers of the English Dominican Province.
His starting point is the term ‘being’. He assumes that being is the first object of apprehension. Whenever human beings apprehend something they know that it is a being because it is included in all things. Therefore, for whoever apprehends the term being, it is self-evident that "the same thing cannot be affirmed and denied at the same time," (Principle of non-contradiction (PNC). Aristotle, *Metaphysics* 4. 9.) This is the first principle of theoretical reasoning and on this principle all others are based. Likewise, "good" is the first object of the apprehension of the practical reason. Following Aristotle, Thomas claims that since every agent acts for an end under the aspect of good, it is the ground of the first principle of practical reason that “good is that which all things seek after” (Aristotle, *Nicomachean Ethics* 1.1). Given these premises, therefore, the first precept of natural law is “seek good and avoid evil.” This is the basic principle from which all other precepts of the natural law are derived. So that whatever the practical reason naturally apprehends as man’s good (or evil) belongs to the precepts of the natural law as something to be done or avoided.

Because of the manifold nature of human beings; as a substance, as an animal, and as a rational being, this basic principle becomes manifest in three formulations. First, the way human beings incline to seek good as a substance is to preserve their own being. So, the means of preserving human life, and of warding off its obstacles, belongs to the natural law. Second, as an animal, human beings incline naturally to preserve their own species through procreation and raising their offspring. Actions regarding these belong to the natural law. Finally, as a rational being, human beings incline to know the truth, and to live in a society. So, they have to shun ignorance, to avoid offending those among whom one has to live, etc.
Thomas provides a rationale for the first principles of human action, which are apprehended by human beings without reasoning, by virtue of their very nature. This is also an intellectualist answer to the question why should conscience be followed?, as opposed to Bonaventure’s voluntaristic answer. Hence “synderesis is said to be the law of our mind, because it is a disposition containing the precepts of the natural law, which are the first principles of human actions” (ST 2 – 1.94.1c). This argument also tries to show that these precepts are necessarily true because they are derived from self-evident truths and hence can serve as the major premises for further reasoning. In this way, Thomas also guarantees the infallibility of synderesis.

In replies to the arguments that synderesis is a potentiality, Thomas clarifies its nature further. Unlike his predecessors, he does not assign a single place to synderesis within reason but argues that it is shared by both lower and higher reason. He gives an argument for this controversial claim, because it is assumed that if a disposition is pertinent to the unchangeable truths, then it belongs to the higher reason. However, Thomas elucidates the relation among synderesis, higher reason, and unchangeable things. He states that there are two kinds of unchangeable things. First are the objects of higher reason that are divine things, whose nature do not change. Second are the necessarily true propositions. For example, “every whole is greater than any of its parts.” Although this proposition is applied to the changeable things the relation between whole and its parts is necessarily true and hence unchangeable. Actually, Thomas distinguishes

44 There is no doubt that PNC is necessarily true, but the second principle, “good is that which all things seek after” is at least open to discussion. Some argue that sometimes human beings desire bad for its own sake. See, Michael Stocker, “Desiring the Bad: An Essay in Moral Psychology”, The Journal of Philosophy, Vol. 76, No. 12 (Dec., 1979), pp. 738-753.
ontological and logical senses of being unchangeable. The ontological sense requires that the being stays the same, whereas logical sense requires the relation between beings stays the same. Given this explanation it is obvious that “synderesis aims at what is unchangeable” in the logical sense and hence it should not be in the higher reason alone but it is shared by both higher and lower reason. This conclusion also supports the infallibility of synderesis (De verit., q.16, a.1, ad 9, p.125).

Although Thomas establishes the infallibility of synderesis, he offers further arguments to support this claim in his treatment of the second question: Can synderesis do wrong? (De verit., q.16, a.2. p. 127) His answer is justified by a foundationalist theory of knowledge. Following the authority of Aristotle (Physics 1.6. 189a19), he states that the first principles are always “permanent and unchangeable and conserve the right order” in all the works of nature. So that anything changeable can be traced back to the first principles, which constitute the foundation of our knowledge. Similarly, “every particular apprehension comes from some absolutely certain apprehension about which there is no mistake. This is apprehension of basic principles, by reference to which all particular apprehension are tested and in virtue of everything true wins approval but everything false is rejected” (De verit., q.16, a.2c., p. 127). The same is true about the rightness of human actions. There are some unchangeable principles by reference to which rightness of the deeds are tested and only the ones which pass the test are good. Since these principles constitute synderesis it can never do wrong but always murmurs back in reply to evil and turns humans towards what is good.
In replies to the argument that *synderesis* can do wrong, Thomas examines a case where it seems to be wrong. People who read John 14:2 “The time comes when everyone who kills you will judge that he does God a service”, interpret that to kill the apostles will be a service to God. Hence, by the actualization of *synderesis* they sin. Thomas remarks that the basic principle in *synderesis* is “Serve God” and there is no mistake in it. For *anyone* who understands the term God knows that the proposition “Serve God” is necessarily true. However, killing the apostles will be a service to God is a false judgment of reason. Therefore, sin is not an outcome of the actualization of *synderesis* but erroneous reasoning.

As expected, the next question will be, “is *synderesis* extinguished in some people?” (*De verit.*, q.16, a.3, p. 128). For, if its job is to murmur back in reply to evil and to turn humans towards what is good, then why it does not function in cases like the one above? To answer this question, Thomas distinguishes two functions of *synderesis*. By nature humans know the basic principles of theoretical and practical reasoning and *synderesis* is the innate disposition by which humans know the basic principles of practical reasoning. This is the first function of *synderesis*, i.e. to contain all the basic principles. Second, it serves as the major premise in the deductive reasoning. This function is “to turn humans towards what is good” because the major premise shows the way to be followed. Now it is the job of reasoning to follow this general direction and draw particular conclusions. In this process reason may be “swallowed up” by appetitive powers and make mistakes. Obviously, this does not show that *synderesis* is extinguished in some people because it was there in the right place in the syllogism but

---

45 Thomas appropriates Aristotle’s examination of *akrasia* here.
reason fails to draw the right conclusion from it. Thomas identifies one more condition other than intervention of appetitive powers that is the injury of the organs. When the organs are damaged they cannot provide reason the required information for the right judgment. In this case it is not possible to exercise free choice or to reason properly and hence to derive conclusions. Thomas’ consideration of physical deficiencies is a novelty in the analysis of the reasons of the obstructions of the operation of syndereisis.

However, the “murmuring back in reply to evil” part of syndereisis’ function is missing in this analysis. Thomas does not say anything about why syndereisis does not murmur back in reply to a wrong conclusion. Examination of his treatment of conscientia may reveal his answer to this question.

Thomas’ analysis of conscientia starts with a classificatory question: “Is conscientia a potentiality, a disposition or an actualization?” (De verit., q.17, a.1, p.130). Formulation of the question foreshadows the answer. Being an actualization is not considered to be an option for conscientia before Thomas. His classification is based on its being an application of knowledge to particular cases. He justifies this claim with an analysis of the different meanings of conscientia. ‘Conscientia’ is used for:

1) the thing of which one is conscious, as ‘belief’ is used for the thing believed;
2) the potentiality or disposition by which we are conscious;
3) the actualization (De verit., q.17, a.1c, p.130).

Thomas observes that sometimes a single noun is used for a potentiality, a disposition, an object and an actualization. Following Aristotle’s suggestion (Topics 2.2.110a16), he says that “words should be used as most people use them”. Since, in everyday language
conscientia is used in the first sense, it should be taken in this sense. He remarks that this sense cannot be used for a potentiality or a disposition but only for an actualization. In addition, only this sense covers everything about conscientia.

Thomas gives an argument to support this observation. He claims that the same name can be used for a potentiality, a disposition, and an actualization only when there is one way of actualizing the potentiality or disposition in question, “as seeing is proper to the visual potentiality, and knowing is actualization of the disposition of knowledge: thus ‘sight’ sometimes means potentiality, sometimes its actualization, and ‘knowledge’ similarly” (De verit., q.17, a.1c, p.130). However, most of the time names of the actualizations can be associated with more than one potentiality or disposition, as in the use of the word ‘use’. It may signify the actualization of any disposition and potentiality. Thomas argues that this is also true for ‘conscientia’ because it “signifies the application of knowledge to something, so that to be conscious of something (conscire)” (De verit., q.17, a.1c, p.130). Since any knowledge can be applied to something, it cannot be the name of some special disposition, or some potentiality. Hence, conscientia is the name of “the actualization, which is the application of some disposition or other, or of something or other known, to a particular actualization” (De verit., q.17, a.1c, p.131)

At first sight, this argument seems to be a weak one because conscientia is assumed not to be the application of any knowledge but a specific one, namely synderesis. So, it seems to be possible that there is one way of actualizing synderesis, and the name of the actualization can be synderesis as well, or conscientia can be used of instead of synderesis and signify both the disposition and its actualization, i.e., of having
some basic general principles and their application. However, Thomas is quite consistent in his account as can be seen below.

Thomas first distinguishes two ways of applying something known to a particular actualization:

In the first we consider whether there is or was an actualization...we are said to have consciousness of an actualization that has occurred or not occurred. E.g. when, in everyday usage, it is said: ‘I have no consciousness that this occurred’, i.e., I do or did not know whether this has occurred or did occur. In this sense that Genesis 43:22, ‘We have no consciousness of who put money in our sacks’, and Ecclesiastes 7:23, ‘Your conscience knows that many times you have yourself cursed others’, to be understood. And in this sense that conscientia is said to be bear witness to something, e.g., ‘My conscience bears me witness’ (Romans 9:1) (De verit., q.17, a.1c, p.131)

In this way, conscientia is an awareness of one’s possessing of any kind of knowledge. My conscientia bears me witness that I know so and so. For this reason Thomas claims that conscientia is an actualization (as being conscious of) of knowledge (as a disposition). Actually, this is the literal meaning of conscientia, and Thomas is right when he says that it can be the application of any kind knowledge.

The second way of applying something known to a particular actualization is to check whether the actualization is right or not and there are two forms of it: “In one we are directed through the disposition of knowledge to do or not to do something. In the other, the actualization is tested, after it has taken place, by disposition of knowledge, for whether it is right or not right” (De verit., q.17, a.1c, p.131). Thomas states that these two forms of application correspond to the two methods in theoretical matters, namely, discovery (induction) and judgments (deduction). By “tracking down conclusions from
premises”, we *discover* the knowledge about what should be done. By tracing conclusions back to premises, we *judge* whether what has already been done is right or not.

Conscientia takes both forms:

When knowledge is applied to an actualization in order to direct it, conscientia is said to goad or urge or bind us. But when knowledge is applied to an actualization by way of testing what has already been done, conscientia is said to accuse or worry us if what has occurred is found to be out of accord with the knowledge by which it was tested, and to defend or excuse us if what has occurred is found to have turned out in accordance with the piece of knowledge (De verit., q.17, a.1c, p.131).

Thomas specifies two functions of *conscientia* with this distinction; one answers the question ‘what should I do’ prospectively, the other answers the question ‘did I do the right thing’ retrospectively. In both cases, disposition of *synderesis*, the disposition of wisdom, and the disposition of knowledge (whether all together, or just one of them) are applied to the actualization.\(^{46}\) Hence, *conscientia* is the actualization of these dispositions. In this sense *conscientia* is not simply an application of the general principles to the particular cases or derivation of particular conclusions from basic principles, but it is to apply different kinds of knowledge to produce rules to direct and judge one’s ethical behavior, e.g. to use *synderesis* as guiding principles, wisdom to derive particular conclusions and knowledge of particulars to combine them together.\(^{47}\)

---

\(^{46}\) In the first method of application, by which knowledge is applied to an actualization to know whether it occurred, is an application to a particular actualization of perceptual information, e.g. (by which we recall what was done) or of perception, by which we perceive this particular actualization that we now do (De verit., q.17, a.2c, pp. 131-2).

\(^{47}\) Sorabji limits the function of *conscientia* only to the application of *synderesis*: “*synderesis* is a disposition disposing us to values so universal as to ensure its infallibility, and *conscientia* is a fallible act of applying *synderesis*” (Sorabji 2014, p.64).
This clarification set the stage for the next question: Can conscientia be mistaken? For, it seems to be possible to make mistakes in the process of application of knowledge to formulate rules. According to Thomas, mistakes are due to two reasons: either what is applied contains a mistake or it is not applied properly. Similarly, mistakes in reasoning can occur in two ways: either there is a false premise or reason is not used correctly.

In the first case, since the proposition in the synderesis is too general, it cannot be applied without a particular premise. Once this particular premise is provided, conscientia can be completed. As it is discussed above, the judgments of synderesis are infallible, but higher reason may make a mistake when it mistakenly identifies something in accordance with the natural law, or against it, which is not, “e.g. heretics who believe that taking an oath has been forbidden by God” (De verit., q.17, a.2c, p.132). Similarly, a mistake in the lower part of reason may cause a mistake in conscientia, by providing wrong information about social norms. For example, “theft, although it is expressly contrary to the natural law, was not considered wrong among the Germans, as Julius Caesar relates (De Bello Gall. vi)” (ST 2 -1.94.4c).

In the second case, as a result of a failure of using a valid form of argument, a false conclusion is drawn, in the process of application of knowledge. Because of this type of faulty reasoning conscientia can be mistaken. However, Thomas argues that in certain cases conscientia can never make a mistake, i.e., when there is a general judgment in synderesis about the particular actualization to which conscientia applied. Since Thomas presumes that there are certain things that we are “required to know,” error regarding these matters would be inexcusable. P. Hannon observes that Aquinas arrives at
this view because he “found it difficult to think that one might bona fide be mistaken about or ignorant of the moral law” (Hannon 2004, p.55n3). Hence, Thomas concludes that conscientia cannot err with regard to propositions such as “I ought to love God” or “evil should not be done,” by thinking that the opposite is acceptable, since general principles of synderesis are directly applied to the particular circumstance in these instances, such that both the major and minor premises are self-evident, given that the minor premise contains ideas from the major premise, which are even expressed in the same terms (De verit., q.17, a.2c, p.132). However, people can make mistakes even in identifying a particular instance of a universal term. Hence, his argument about the cases where conscientia can never make mistake, is not a convincing one.

In the next question, Thomas discusses whether the conscientia binds (De verit., q.17, a.3, p.134). He believes that it binds but to lay out how it binds, first he examines the meaning of ‘binding’. He claims that ‘binding’ is used metaphorically for spiritual things to imply necessity, which is among physical beings. He explains that:

There are two kinds of necessity which can be imposed by another agent. The first is a necessity of force, through which everything absolutely necessarily has to do what is determined by the action of the agent; the other should not strictly be called force but rather inducement. This is a conditional necessity that is, derived from a goal; e.g. there may be a necessity imposed upon someone that, if he does not do such-and-such, he will not obtain his reward.

The first kind of necessity, which is that of force, does not occur in changes of the will, but only in bodily things, because the will is necessarily free from force. The second kind of necessity can be imposed upon the will, e.g. it may be necessary to choose such-and-such, if a certain good is to result, or if a certain evil is to be avoided...But just as the necessity of force is imposed on bodily things by some action, so conditional necessity is imposed upon the will by some action. The action, by which the will is changed, however, is the command of a ruler or the governor...Thus the command of something which governs is related to binding bodily things by necessity of force. But the action of a bodily agent only introduces necessity into another thing on which it acts; so
someone is only bound by the command of a ruler or lord, too, if the command reaches him who is
commanded; and it reaches him through knowledge (De verit., q.17, a.3c, p.134).

Thomas specifies the conditions, which make conscientia binding on the basis of the
analogy between force and injunction. Hence, conscientia is binding because it aims to
seek good and avoid evil, which is the command of an absolute ruler, i.e. God. So, it is
necessary to follow conscientia to obtain the reward, i.e., salvation. However, there are
certain conditions, which makes conscientia binding. First of all, the agent has to be
capable of being informed of the command. Second, the agent has to know the command
because it may be the case that he is ignorant of it. This is a legitimate excuse unless the
command in question is one of those one is obliged to know.

Thomas brings new insight to the discussion in the subsequent question: Does a
mistaken conscientia bind? He says that a correct conscientia binds without qualification
and per se, whereas, a mistaken conscientia binds relatively and accidentally. The former
is obvious given the discussion of the previous question, for it is the injunction of god.
Although, a mistaken conscientia does not have such strength, it is still binding,
according to Thomas, because it is believed to be correct. He states that if something is
desired or pursued for its own sake then it is pursued per se but if it is desired or pursued
for the sake of something else, then it is pursued accidentally. For example, if someone
loves wine for its sweetness, then he loves it accidentally. Similarly, a mistaken
conscientia is binding because one follows it for the sake of fulfilling the command of
God, by doing the right thing on the basis of seeking good and avoiding evil, which
happens to be mistaken. For this reason, a mistaken conscientia is binding accidentally.
Potts claims that Aquinas’ argument that a mistaken *conscientia* is binding is not a compelling one on the ground that:

Aquinas maintains that a mistaken conscience is binding on the ground that ‘We ought to obey God’s commands is a basic deontic proposition; if it is not, then at least some alternative meta-rule must be shown to be basic in order to maintain that a mistaken conscience can ever excuse. Otherwise, we have not the slightest ground to suppose that there is a valid consequence from “a believes that he ought to φ” to a ought to φ”, anymore than from “ A believes that p” to “p” (Potts 1981, p.60).

I agree with D’Arcy that actually, Aquinas has such a meta-rule that is “seek good and shun evil”. According to him Aquinas views it is as a purely formal principle that governs all our moral reasoning (D’Arcy 1961, 52). As shown above, Aquinas provided an argument for this meta-rule, which can stand alone without the presupposition that “We ought to obey God’s commands”.

Still, Thomas adds a new component to the understanding of *conscientia*; one’s intention to do the right thing is part of *conscientia*. Bonaventure thinks that desiring to do the right thing is a part of *synderesis*, therefore, as a motivation it is in the appetite part of the soul. However, Thomas thinks that having good faith is part of *conscientia* because *conscientia* is the means of obtaining reward, i.e. approval of God, whose first requirement is to have faith.

Thomas’ treatment of conscience brings together and clarifies the previous thinkers’ views on conscience. He followed Philips’s path by placing *synderesis* and *conscientia* in the apprehensive part of the soul and providing cognitive explanations about their function. He concludes that synderesis is an innate non-deliberative inclination toward good. He further holds that this innate capacity nonetheless needs to
acquire material from the senses in order to articulate its terms. This is a useful clarification, as it draws us away from Philip’s immanentism or Bonaventure’s illuminationism. He also removes the drawbacks in Bonaventure’s analysis of the basic general propositions by introducing a criterion for necessary truths.

2.6 Conclusion

This chapter attempts to show that medieval thinkers make substantial contributions to the understanding of conscience while working on the Scripture and the writings of Paul, by appropriating Greco-Roman philosophical ideas. Their treatises on conscience can be read as the attempts to justify the understanding of conscience in the following text by Paul:

For whenever gentiles who do not have the law by nature do what the law requires, they are, without having the law, a law for themselves: these people show the requirements of the law to be written in their hearts while their συνείδησις bears witness and their reasoning in mutual debate accuse or also excuse them, on the day when God judges the secret things of the people, according to my gospel through Christ Jesus (Rom. 2, 12-16).

They try to give an account for the “law written in the heart”, which is supposed to guide human beings to distinguish between right and wrong and “the natural tribunal”, which decides whether they followed this guide. Paul associates conscience with the natural law but the medieval thinkers equate conscience with the natural law. This equation has some consequences: on the one hand conscience is the container of the God given infallible natural law, on the other hand conscience is the fallible judgment about right and wrong. They try to overcome this problem by introducing two aspects of conscience: synderesis and conscientia.
In general, *synderesis* is defined by the medieval thinkers as some kind of potentiality or disposition of general principles, or of the natural law, which constitute the God given infallible content of conscience. However, it is remarkable that regarding the content of *synderesis*, we do not have enough details. (D’Arcy 1961, p.49). Although occasionally some medieval thinkers provide additional examples of specific content, the most quoted basic principle is: Seek good; shun evil. Even though, Thomas claims that all the precepts of the natural law are the universal first principles of *synderesis*, still the content of *synderesis* is highly abstract and needs to be specified. Another problem, medieval thinkers raise and try to solve regarding the content of *synderesis* is that the nature of the content: whether it contains some general propositions (Philip) or some necessary truths (Bonaventure and Thomas), and if so what is the criterion to determine whether they are general or necessary? Although their accounts are different and none of them is satisfactory, still their attempt contributes not only to enriching the understanding of conscience but to philosophy in general. Especially, their discussion about the nature of general propositions set the ground for the further clarifications in ethics and epistemology.

*Conscientia* is the fallible part of conscience that is responsible for drawing particular conclusions from *synderesis* with the help of practical reason. By the time of Thomas it is described as an act of judgment of practical reason, which is the conclusion of the process of an application of universal moral principles to the particular situation. Thomas presented this in terms of a syllogism. Although this marked a definitive stage in the understanding of conscience, it was not completely unrelated to what had gone before, as even Philip had already alluded to some process of application being in
operation. This process is subject to flaw and failure, but nevertheless, the medievals regard it as binding.

Medieval thinkers develop a model to explain the workings of conscience by adopting Aristotle’s account of practical reasoning where the major premise is given by *nous* or moral virtue, the minor premise is given by *phronesis*. In the medieval model the major premise is provided by *synderesis*, and the minor premise is still given by the practical reason. As a novelty, they introduce *conscientia* to draw the conclusion. Any failure in this process is explained following Aristotle’s explanations of *akrasia*.

According to the medieval thinkers, the function of *synderesis* is to incline human beings towards good on the basis of the natural law and to murmur back when they go astray. As we have seen above, they provide detailed accounts how it inclines us towards good but they do not explain the mechanism how *synderesis* becomes aware of wrongdoings. Since *synderesis* is not capable of judgment, it depends on the judgments of *conscientia* and when *conscientia* errs, *synderesis* remain ineffective. This deficiency may be due to their motivation to keep *synderesis* away from the possible errors of judgment, so that it remains infallible.

Medieval thinkers’ major contribution to the understanding of conscience is through the questions they raise about the content and function of conscience. Their answers provide the ground for further clarifications as we will see in the next chapter.
CHAPTER III
Kant

3.1 Introduction

There are a few depictions of conscience in Kant’s writings. An extended treatment of conscience may be found in Religion within the Boundaries of Mere Reason (Religion, 1793) but a full account of conscience appears in the Metaphysics of Morals (MM, 1797). Since the concept appears rarely in Kant’s works and the full treatment comes only after his two well-known books in ethics, the Groundwork of the Metaphysics of Morals (GMM, 1785) and the Critique of Practical Reason (CPrR, 1788), scholars do not pay much attention to the subject. The existing literature on Kant’s theory of conscience examines the role of conscience within his ethics, i.e., whether it plays a crucial role in it, if so, to what extent. Although all of these scholars provide valuable insights into the

---


49 Hegel provides a phenomenological critique of Kant’s theory of conscience in the “Conscience. The ‘Beautiful Soul’, Evil and its Forgiveness” section of the Phenomenology of Spirit. This is a topic my dissertation cannot engage due to time and space constraints. In another work I wish to discuss the Kantian understanding of conscience with reference to Hegel, Schelling and Fichte. Thomas Hill gives a detailed analysis of Kant’s understanding of conscience in “Four Conceptions of Conscience”, in Hill, Human Welfare and Moral Worth: Kantian Perspectives (Oxford: Clarendon Press, 2002). Allen Wood also has included a lengthy discussion of conscience in Allen W. Wood, Kantian Ethics (New York: Cambridge University Press, 2008). Korsgaard and Barbara Herman both mention conscience within their discussions of more familiar aspects of Kant’s theory, but neither of them explores Kant’s own comments on conscience. On the theme of conscience and Kant one could also consult J. David Velleman’s “The Voice of Conscience”, Proceedings of the Aristotelian Society 99 (1998), pp. 57-76. In Velleman’s case, however, the focus is not on what Kant wrote about conscience, but rather on how we can think of the more familiar Kantian apparatus (duty, moral law) in terms of conscience. Dean Moyar, in a recent article ("Unstable Autonomy: Conscience and Judgment in Kant’s Moral Philosophy", in Journal of Moral Philosophy 5 (2008), pp. 327–360), provides a detailed analysis of the concept and claims that Kant’s reflections on conscience led him towards a Fichtean view of the primacy of a pure act of self-consciousness and the irreducibility of the individual’s authority in moral judgment to the universal law.
understanding of Kant’s theory of conscience, none of them refers to Kant’s contribution to the understanding of the concept, which will be the main aim of this chapter. I argue that Kant’s account of conscience can be interpreted as a rehabilitation of medieval understanding of the subject. Although there is no direct reference to the medieval thinkers, it can be claimed that in his works Kant tries to improve their understanding of conscience. In his ethics lectures, Kant uses Baumgarten’s text books, where there is a chapter on conscience. In these lectures, Kant evaluates Baumgarten’s views on conscience, which are not very different from the medieval understanding and then gives his own account of conscience. In his published writings, Kant basically reiterates the points he made in his lectures in a more eloquent style.

In order to specify the ground of Kant’s examination of conscience, I give a brief account of his critical philosophy. The distinction between pre-critical and critical designates the difference in Kant’s approach between before and after the publication of *Critique of Pure Reason* in 1781. Kant defines his critical approach in the following way: “…reason should take on anew, the most difficult of all its tasks, namely that of self-knowledge, and to institute a court of justice, by which reason may secure its rightful claims while dismissing all its groundless pretensions, and this is not by mere decrees but according to *its own* eternal and unchangeable laws…” (A XII, emphasis added). By assigning and carrying out this task that is to determine the limits and status of the claims reason can make, Kant would be able to show the role of human cognition in the knowledge acquisition. He revolutionizes philosophy in a Copernican sense by changing the setting of the inquiry. He claims that:
up to now it has been assumed that all our cognition must conform to the objects; but all attempts to find out something about them a priori through concepts that would extend our cognition have on this presupposition, come to nothing. Hence let us once try whether we do not get farther with the problems of metaphysics by assuming that the objects must conform to our cognition, which would agree better with the requested possibility of an a priori cognition of them, which is to establish something about objects before they are given to us (Kant 1787, *CPR*, BXVI, emphasis added).

One of the outcomes of this assumption is that “our representation of things as they are given to us, does not conform to these things as they are in themselves, but rather that these objects, as appearances conform to our manner of representation” (Kant 1787, *CPR*, BXX). Hence, on the one hand, Kant guarantees the validity of the foundational principles for the appearances/phenomena, on the other hand he leaves open the possibility that the things in themselves/ noumena may not be subject to these principles. The distinction between phenomena and noumena enables Kant to have room for faith and morality. While as phenomena we are bound by the deterministic laws of nature, as noumena we are free to bind ourselves to the moral laws that are formulated by our practical reason.

Kant’s understanding of conscience will be evaluated within the framework of his critical philosophy and his pre-critical writings and lectures on ethics will not be taken into account in this study.\(^5^0\) The focus of this chapter will be on Kant’s final account of conscience in the *Metaphysics of Morals* (*MM*), however, his other critical writings on conscience will be taken into account as well. In the *MM*, Kant treats conscience under

two headings: (1) as one of the moral feelings presupposed by the mind’s receptivity to the concept of duty (6:401), and (2) as the duty of self-examination and self-judgment by our own internal judge (6: 438-41). The second one is Kant’s primary account of conscience; however, it will be helpful to start with the first one to understand the nature of conscience.

3.2 Conscience as a Moral Feeling

Before we examine Kant’s account of conscience as a feeling, it is crucial to clarify the place of feelings in his ethics. Definitely, Kant’s Copernican revolution in philosophy effects his understanding of the feelings but this understanding is misunderstood commonly. Allen Wood explicates this misunderstanding in the following way:

…action with genuine moral worth must be unaccompanied by either feeling or desire. On the contrary, Kant’s psychology of action involves the thesis that all action involves the representation of an end to be produced, and a desire for that end, and the conception of desire for an end as the representation of it accompanied by a feeling of pleasure (if the feeling is displeasure, then it is a case not of desire but of aversion). The key point to understand is that for Kant, not all desire is inclination. ‘Inclination’ refers only to habitual empirical desire. Fundamental to Kant’s psychology of action … is the idea that not all desire arises passively through our receptiveness to empirical impulses, but some desires can result solely from the activity of pure reason, the free and self-directing side of our nature…In sensible creatures such as human beings, purely rational desires, like empirical desires, also manifest themselves in the form of feelings – feelings resulting directly from the operation of reason on our sensibility. Susceptibility to these feelings is essential to our capacity to act rationally, and a being who was not susceptible to them could not be a responsible moral agent at all (Wood 2009, p.3).

The opening sentences of the first section on conscience entitled “Concepts of What is Presupposed on the Part of Feeling by the Mind’s Receptivity to Concepts of Duty as such (Ästhetische Vorbegriffe der Empfänglichkeit des Gemüts für Pflichtbegriffe überhaupt)”, supports Wood’s claims about the role of feelings. Kant writes:
There are certain moral endowments such that anyone lacking them could have no duty to acquire them. They are moral feeling, conscience, love of one’s neighbor, and respect for oneself (self-esteem). There is no obligation to have these because they lie at the basis of morality, as subjective conditions of receptiveness to the concept of duty, not as objective conditions of morality. All of them are natural predispositions of the mind (praedispositio) for being affected by concepts of duty, antecedent predispositions on the side of feeling. To have these predispositions cannot be considered a duty; rather, every human being has them, and it is by virtue of them that he can be put under obligation. Consciousness of them is not of empirical origin; it can, instead, only follow from consciousness of a moral law, as the effect this has on the mind (Kant 1797, MM. Ak. 6:399).

It is remarkable that Kant conceives conscience as a natural predisposition of the mind.

This is almost the same terminology, medieval thinkers use to describe synderesis, i.e., an innate disposition of reason. Kant also provides a similar explanation for this description; because of our rational nature, these “sensuous pre-concepts” (Ästhetische Vorbegriffe) are part of our natural endowment. They are the preconditions of being a moral agent, only through which it is possible to put human beings under obligation.

Kant’s treatment of the first two of these “moral endowments”; moral feeling and conscience, could be read as a revision of the medieval understanding of synderesis. He clarifies the content and the functions of synderesis and redistributes them between moral feeling and conscience.

Kant starts his analysis with moral feeling that is “the susceptibility to feel pleasure or displeasure merely from being aware that our actions are consistent with or contrary to the law of duty” (Kant 1797, MM. Ak. 6:399). He remarks that “every determination of choice proceeds from the representation of a possible action to the deed through the feeling of pleasure or displeasure” (Kant 1797, MM. Ak. 6:399). It is not a novelty to identify the feeling of pleasure or displeasure with the motivation of action. There are many champions of the claim that human beings seek pleasure and avoid pain
by nature, throughout the history. However, Kant associates pleasure with our awareness that our actions are consistent with the law of duty. We feel pleasure when we do the right thing or think about doing it, which is nothing but to act in accordance with the law of duty or we feel displeasure whenever we act contrary to it.\(^{51}\) The concept of moral feeling is very similar to the conceptualization of *synderesis* as a disposition of reason that is always directed to the good. However, medieval thinkers assume that human beings are directed towards good by nature because again by nature, they possess the knowledge of the good. Actually, this is one of the pitfalls of the medieval understanding of *synderesis* that is to provide a convincing account for how human beings possess the knowledge of the good, as we have seen in the previous chapter.

Since Kant is very well aware of the difficulty of justification of the knowledge of the good, he warns that “it is inappropriate to call this feeling a moral *sense*, for by the word ‘sense’ is usually understood a theoretical capacity for perception, directed toward an object, whereas moral feeling (like pleasure and displeasure in general) is something merely subjective, which yields no cognition” (Kant 1797, *MM. Ak*. 6:400). Moreover, he insists that human beings do not have “a special ‘sense’ for what is good and evil”. Rather they have “a *susceptibility* on the part of free choice to be moved by pure practical reason (and its law), and this is what we call moral feeling” (Kant 1797, *MM. Ak*. 6:400).

---

\(^{51}\) In the *GMM*, Kant criticizes moral feeling explicitly for not being “fit to be the ground of moral laws” (Ak. 4:442). Nevertheless, he admits the motivational force of the moral feelings when they originate from reason, almost with the same words a couple pages later in the *GMM*: “In order for a sensibly affected rational being to will that for which reason alone prescribes the ‘ought,’ it is admittedly required that his reason have the capacity to induce a *feeling of pleasure* or of delight in the fulfillment of the duty...” (Ak. 4: 460).
According to Kant, it is not possible for human beings to be without moral feeling. If they lacked this feeling they would be considered as morally dead and there would be no difference between human beings and the rest of natural beings. There is no duty to have moral feeling or to acquire it because “every human being (as a moral being) has it in him originally” (Kant 1797, MM. Ak. 6:399). Like synderesis, moral feeling is built in our nature and it cannot be extinguished.\(^5^2\)

Kant starts his description of conscience with exactly the same sentence he uses for moral feeling: “So too, conscience is not something that can be acquired, and we have no duty to provide ourselves with one; rather, every human being, as a moral being, has a conscience within him originally” (Kant 1797, MM. Ak. 6:400). To have a conscience cannot be a duty itself, if it is, then it would be equal to having a duty to recognize duties and this may go \textit{ad infinitum}. Therefore, Kant claims that

conscience is practical reason holding the human being’s duty before him for his acquittal or condemnation in every case that comes under a law. Thus it is not directed to an object but merely to the subject (\textit{to affect moral feeling by its act}), and so it is not something incumbent on one, a duty, but rather an unavoidable fact (Kant 1797, MM. Ak. 6:400).

Conscience is the recognition of duty as to act in a certain way or to refrain from action and moral feeling is the confirmation of conscience’s verdict by feeling pleasure or displeasure about the action and in this way conscience affects moral feeling. Hence, it

\(^5^2\) In \textit{Religion}, Kant explains this predisposition towards good, following the medieval line: “For, in spite of that \textit{fall}, the command that we ought to become better human beings still resounds unabated in our souls...[because] there is still a \textit{germ of goodness} left in its entire purity, a germ that cannot be \textit{extirpated} or \textit{corrupted}...[This] original good is holiness of maxims in the compliance to one’s duty...” (Kant 1793, \textit{Religion. Ak.} 6:45-46, emphasis added).
can be claimed that Kant redistributes the functions of *synderesis* between moral feeling and conscience:

1. to turn towards the good to the moral feeling (=feeling pleasure when our actions are consistent with the law of duty)

2. to recognize what is good to conscience (=conceiving whether an action is consistent with the law of duty).

Kant’s identification of conscience as a sensuous pre-concept is remarkable. In the Collins’ lecture notes; it is even identified as an instinct:

> It is not a mere faculty, but an *instinct*, not to pass judgment on, but to direct oneself. We have a faculty of judging ourselves according to moral laws. But out of this we can make use as we please. Conscience, however, has a driving force, to summon us against our will before the judgment-seat, in regard to the lawfulness of our actions. It is thus an instinct, and not merely a faculty of judgment. Moreover, it is an instinct to direct and not to judge (Kant.1785, *Collins. AK*. 27:351).

In the rest of the lecture he dwells on the difference between “to direct” and “to judge” by appealing to the courtroom metaphor for the first time, which is going to be a consistent trait of his discussions on conscience starting with this lecture. I will develop a detailed analysis of this metaphor in this chapter.

Kant states that conscience is an unavoidable act of judging actions and argues that it is not possible to have no conscience at all because in such a case, a “human being could not even conceive of the duty to have one, since he would neither impute anything to himself as conforming with duty nor reproach himself with anything contrary to duty” (Kant 1797, *MM. Ak.* 6:401). For this reason, conscience is a sensuous pre-concept that makes the conception of duty possible, without which deeds cannot be imputable to
someone. A few pages later, in his treatment of freedom, Kant even writes that
“conscience was treated earlier as the condition of all duties as such” (Kant 1797, MM. Ak. 6:407). So, when we talk about someone who does not have conscience, it can only mean that “he pays no heed to its verdict” (Kant. Ak. 6:401).

Before moving to the courtroom analysis, I would like to highlight the importance of Kant’s identification of conscience with a feeling and/or instinct. He does not insist on this identification because it is very hard to capture the real nature of conscience. Most of the thinkers we have studied so far identified it with “a part of reason” but still none of them are able to tell exactly its difference from reason. Conscience seems to be one of the abilities of reason but this explanation is not satisfactory because it is more than the conclusions of practical reason. Some of the thinkers associate it with will and with some of the emotions like guilt and remorse. However, these accounts are not convincing either because these emotions seem to depend on moral judgments and there is a gap between

53 This remark recalls the opening sentence of the GMM, where Kant writes: “It is impossible to think of anything at all in the world, or indeed even beyond it, that could be considered good without limitation except a good will.” (Kant 1785. Ak. 4: 393). Both conscience and good will can be defined with reference to duty. Conscience is the recognition of the duty, whereas good will is the volition of the duty.

54 In Vigilatius lectures, Kant explains the case of these type of persons: “Consciousness must be accompanied with an attitude of sincerity, i.e., that the subject be aware of having entered upon his examination with an eye to probability; this examination always has to do, of course, with the merely external circumstances in the action; it calls for a customary rigor, in order not to view a factum as other than it really is; man is only too readily inclined to persuade himself of something, and conjure up more than the truth. There are tendencies, indeed, in the souls of many, to make no rigorous judgment of themselves—an urge to dispense with conscience. If this lack of conscientiousness is already, in fact present, we never get that person to deal honestly with himself. We find in such a people that they are averse to any close investigation of their actions, and shy away from it, endeavoring, on the contrary, to discover subjective grounds on which to find a thing right or wrong” (Kant 1793. Ak. 27: 616).

55 In his Opus Postumum, he refers to this “indefinable” phenomenon even as being (Wesen): “There is a being (Wesen) in me, which is different from me and which stands in an efficient causal relation (nexus effectivus) toward myself; itself free (that is, not being dependent upon the laws of nature in space and time) it judges me inwardly (justifies or condemns); and I, man, am myself this being—it is not a substance outside me. What is most surprising is that this causality is a determination (of my will) to action in freedom [that is], not as a natural necessity (Kant. Ak.21:25).
these emotions and judgments. If there is such an entity as conscience, then its nature is very complicated and very hard to provide a causal explanation for. When we try to give an intentional explanation we have not found the right terms. For this reason Kant’s identification of conscience with a feeling/instinct catches a glimpse of the nature of conscience and when we ask what it is, Kant does not have much to say about what kind of a being it is. Instead he uses the courtroom metaphor to describe it. In this metaphor, conscience plays different roles at the same time, such as being an ability of reason, a kind of feeling, and a sensuous pre-concept at the same time.

3.3. Conscience as an Inner Court of Moral Judgment

In the section “On the Human Being’s Duty to himself as His Own Innate Judge”, Kant gives a detailed account of conscience, where he dwells on the court metaphor. This is not the first use of this metaphor in the history of conscience; even Kant described conscience as a court in his Collins lectures in 1785.56

The inner judicial proceeding of conscience may be aptly compared with an external court of law. Thus we find within us an accuser, who could not exist, however, if there were no law; though the latter is no part of the civil positive law, but resides in reason…In addition, there is also at the same time in the human being an advocate, namely self-love, who excuses him and makes many an objection to the accusation, whereupon the accuser seeks in turn to rebut the objections. Lastly we find in ourselves a judge, who either acquits or condemns us (Kant. Ak. 27:354).

Wood is right when he claims that the description of conscience as a court is not “as metaphorical as we might think” (Wood 2009, p.5). As concluded above, conducting a

56 Describing conscience in legal terms is common in the Hellenistic period (see Chapter 1, for a list), but probably, the first use of the court metaphor for conscience can be found in Philo of Alexandria’ writings. In Flaccus, he introduces the metaphor of the inner court: “He who does wrong knowingly has no excuse, as he is already condemned by the court of conscience (en to syneidotos diskaterion) Flaccum 2.6. Medieval thinkers describe synderesis as a natural tribunal (natural iudicatorium) as well. See also, Chapter 2.
further investigation in regard to the lawfulness of one’s action is one of the functions of conscience. This is carried out with the collaboration of the accuser, the defender and the judge, whose job is very similar to the prosecutor, the positive law, the defense attorney and the judge in an external court. The same idea is revisited in the *MM*, but there is difference in the role of conscience. This time it is not the inner judicial proceeding but “the consciousness of an inner court in the human being”:

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 meritem aut demeritum*), belongs to the faculty 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 judicial proceeding [*Gericht* (*coram iudicio*), which, as a moral person giving effect to a law, is called a court [*Gerichtshof* (*forum*). – Consciousness of an inner court in the human being (“before which his thoughts accuse or excuse one another”) is conscience (Kant 1797, *MM. Ak.* 6:438).

However, in the following paragraphs, Kant claims that conscience is also more than this awareness. For example, he states that “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” (Kant 1797, *MM. Ak.* 6:438). So, either “the internal judge” is the same as conscience or it is one of the functions of conscience. Furthermore, Kant is not very precise in identifying a judge’s duties, which are definitely not to observe and to threaten. In addition, a couple of lines below, Kant describes conscience as an “original intellectual and (since it is the thought of duty) moral predisposition”, whose job is to carry out a trial, where accusation, defense and judging take place within the same person (Kant 1797, *MM. Ak.* 6:438). Again, it is not clear whether conscience is the whole trial process or the one who prosecutes and/or judges.
Another metaphor may help to clarify this ambiguity: the metaphor of court of appeals. If we replace conscience as a court with conscience as a court of appeals, Kant’s understanding of conscience will be described coherently.

The duty of a court of appeal is to review decisions of trial courts for errors. Under its standard of review, an appellate court decides the extent of the deference it would give to the lower court’s decision, based on whether the appeal was one of fact or one of law. When reviewing a lower court’s decisions on an issue of fact, courts of appeal generally look for "clear error", i.e. whether the lower court misapplied the facts or the law.

In the light of this definition, Kant’s descriptions of conscience can be reviewed. Practical understanding determines the law, from which we derive rules that command how we ought to act in specific situations. The faculty of judgment is responsible for the imputation of a deed as a case falling under a law. Finally, reason decides whether this imputation is a rightful one. The whole process of determining the lawfulness of an action is similar to the judgments of a lower court. The task of the faculty of judgment is fulfilled by the prosecutor and the defense attorney. Both the prosecutor and the defense attorney provide evidence or refute the evidence to determine the status of an action. The judge takes up the role of reason and concludes whether the action is a lawful one according to the law. Whenever we engage in moral judgment about our own actions this tribunal takes place. As Hill rightly observes, conscience can “come into play only after one has made, or accepted, a moral judgment” (Hill 2002, p. 299). Therefore, conscience is the “consciousness of an inner court in the human being” (Kant 1797, MM. Ak. 6:438). Since conscience is “consciousness of my will, my disposition to the right…a
consciousness of what duty is” (Kant 1793, *Vigilantius. Ak.* 27:614), and “an instinct to direct oneself according to moral laws” (Kant *Ak.* 27:351), it “summons us against our will before the judgment-seat in regard to lawfulness of our actions” (Kant 1785, *Collins. Ak.* 27:351). Once the first order moral judgment is made, conscience automatically takes the case to the court of appeals, where it can be reviewed diligently. Its role is “not to pass judgment but to direct oneself [according to moral laws]” (Kant 1785, *Collins. Ak.* 27:351). Conscience directs reason to judge itself, “whether it has actually undertaken, with all diligence, that examination of actions (whether they are right or wrong), and it calls upon the human being himself to witness for or against himself whether this has taken place or not” (Kant 1793, *Religion. Ak.* 6:186). Therefore, Kant defines conscience as the “*moral faculty of judgment*, passing judgment itself” (Kant 1793, *Religion. Ak.* 6:186). This is exactly what a court of appeals does, to pass judgment on judgment by reviewing the whole trial process diligently.

Since a court of appeals consists of only judges, when we refer to the court, actually we refer to the judges. For this reason, Kant uses the expressions “the inner court of conscience”, “conscience as the inner judge” or “conscience as the sole judge” interchangeably to refer to the same phenomenon, i.e., to the scrutinizer of the *first order moral judgment* (judgment of the lower court).

---

57 The first order moral judgement transforms a maxim into a duty by determining whether the maxim is in accordance with the Universal Law, i.e., whether it is universalizable or not. If the maxim passes the test then it becomes a command of the reason and hence a duty. So, once reason passes judgment what is the right thing to do, i.e., what is the duty, conscience gets involved into the process.

58 Conscience does not pass judgment in the sense whether an action is right or wrong. This is the duty of reason. Conscience passes judgment on whether reason fulfills its tasks properly or not, on the basis of its respect to duty.

58 See the discussion of Moyar below about the nature of this judgment.
Kant points out that, in the trials of conscience the human being has to think of someone other than himself as the judge of his action; “an actual person or a merely ideal person that reason creates for itself” in order to prevent any conflict of interest (Kant 1797, MM. Ak. 6:438). In the footnote he clarifies the reason for this requirement: 

A human being who accuses and judges himself in conscience must think of a dual personality in himself, a doubled self which, on the one hand, has to stand trembling at the bar of a court that is yet entrusted to him, but which, on the other hand, itself administers the office of judge that it holds by innate authority. This requires clarification, if reason is not to fall into self-contradiction. - I, the prosecutor and yet the accused as well, am the same human being (numero idem). But the human being as the subject of the moral lawgiving which proceeds from the concept of freedom and in which he is subject to a law that he gives himself (homo noumenon) is to be regarded as another (specie diversus) from the human being as a sensible being endowed with reason, though only in practical respect – for there is no theory about the causal relation of the intelligible to the sensible – and this specific difference is that of the faculties (higher and lower) of the human being that characterizes him. The first is the prosecutor, against whom the accused is granted a legal adviser (defense counsel). When the proceedings are concluded the internal judge, as a person having power, pronounces the sentence of happiness or misery, as the moral result of the deed. Our reason cannot pursue further his power (as ruler of the world) in this function; we can only revere his unconditional iubeo or veto (Kant. Ak. 6:438).

The aim of this clarification is to justify the judicial authority of conscience as the final judge of action. The authority of conscience is closely related to the authority of moral law. As noumena, human beings are free from all the necessities of the nature. This freedom enables them to give themselves a moral law. Moral law is formulated on the basis of the objective necessity of rationality; hence it has the nature of a command and has legislative authority in the moral realm. Conscience is the consciousness of this force of moral law. This consciousness gives conscience the power to pronounce a sentence about the actions. Therefore, conscience can claim judicial authority about the lawfulness of an action.
Once Kant establishes the judiciary authority of conscience with this highly packed long footnote, he explains the ground of executive authority of conscience. Although he states that the second person could be an actual person, he thinks that reason requires an ideal person because as “the authorized judge of conscience”, is supposed to “impose all action” (Kant 1797, *MM. Ak.* 6:439). Kant describes the ideal person in the following way:

A person in relation to whom all duties whatsoever are to be regarded as also his commands; for conscience is the inner judge of all free actions. Now since such a moral being must also have all power (in heaven and on earth) in order to give effect to his laws (as is necessarily required for the office of judge), and since such an omnipotent moral being is called God, conscience must be thought of as the subjective principle of being accountable to God for all one’s deeds (Kant 1797, *MM. Ak.* 6:439).

Kant conceives the ideal person in terms of divinity. This is not a novelty in his conception of conscience. In the Collins lectures he identifies conscience with the *forum divinum*. He states that there are two kinds of courts, *forum externum* and *forum internum*. The first one is the *forum humanum*, where the human law rules and the second one is *forum conscientiae*, which is the court of moral law. This *forum internum* is a *forum divinum* because “it judges us by our very dispositions, and we cannot, indeed, form a concept of the *forum divinum* other than that we must pass sentence on ourselves according to our dispositions” (Kant 1784-5, *Mrongovius. Ak.* 27:296). Since these

---

59 Kant postulates the idea of God also in the *Critique of Pure Reason* (B XXX; A 634/ B662; A 813/ B 841; A826/B 854), the *Critique of Practical Reason* (5:137) and the *Critique of the Power of Judgment*, Sections 87 – 90. In the last one he adds a footnote that explains the rationale of this postulation: “This moral argument is not meant to provide any objectively valid proof of the existence of God, nor meant to prove to the doubter that there is a God; rather, it is meant to prove that if his moral thinking is to be consistent, he must include the assumption of this proposition among the maxims of his practical reason.” (Kant 1790, *CPJ. Ak.* 5:451)
dispositions are part of the *homo noumenon*, who is both the subject of law giving and subject to the law he gives himself, the *forum externum humanum* cannot judge by them. According to Kant “a forum is required to exercise compulsion; its judgment has to have the force of law; it should be able to compel the execution of the *consectaria* of the law” (Kant 1784-5, *Mrongovius. Ak.* 27:296). In the case of ethical actions, external courts do not have such an authority because the authority and the proof of the facts must be valid according to the moral law.

Kant does not explain explicitly why the *forum conscientia* is the *forum divinum* in the Collins lectures but claims that the judgment of conscience is similar to the divine with respect to the laws they enforce and the infallibility of the judgment. In *MM*, Kant explicitly states that the judging and prosecuting conscience must be thought of as an omniscient, all-commanding, and omnipotent moral being, namely God. Only such a being can secure conscience’s efficacy. However, it does not follow that such an ideal being actually exists outside oneself. For the idea is given subjectively by practical reason, not objectively by theoretical reason. The idea of God, by analogy, “the lawgiver of all rational beings”, puts human beings under obligation to be accountable to “a holy being (morally lawgiving reason)”. Conceptualizing the ideal person in terms of God provides conscience with the authority to be the final judge of moral action. Therewith, this conceptualization also separates Kant from the medieval tradition. God is no longer an objective aspect of the concept but only a regulative idea, which provides conscience with authority.

Identification of conscience with *forum divinum* and the ideal person with God is based on the principle of regulative use of transcendental ideas (Kant. *CPR*, A644/B672).
A specific function of practical reason, namely conscience, is conceptualized with reference to an idea of an omniscient, all-commanding, and omnipotent God. This idea brings unity to the moral cognition. All the moral judgments are appraised from the perspective of such an ideal being.

Given this account of conscience as the divine inner judge, Moyar legitimately asks “how the functions of conscience for Kant can still be kept separate from first-order judgment and deliberation” (Moyar 2008, p. 346). He claims that since conscience uses all the resources of practical reason, the conclusion of conscience is indistinguishable from the conclusion of deliberation given the argument below:

The question for deliberation is to determine what my maxim of action will be. The question before the court is the subjective question, ‘Do you believe that the action you are about to perform is the right (moral) action?’ If you do not believe it is, then you are warned by your conscience. But how do you assess the soundness of your belief? By the same process that would answer the simple question, ‘What is the right (moral) action?’, namely by considering all the appropriate evidence and arriving at a judgment about the most rational/moral action in these circumstances. The separation of a first-order judgment and a subsequent second-order judgment judging itself begins to seem like a rather desperate assertion of the ideal objectivity of moral judgment. In practice it is a distinction without any real difference (Moyar 2008, p. 346-47).

Moyar concludes that the real difference between conscience and the moral judgment is that conscience is the judgment whereby an action is imputed to the subject.

Although Moyar’s interpretation is correct when he says that conscience is a comprehensive judgment of an action, I believe what Kant meant by conscience is slightly different. Actually, Kant repeatedly claims that conscience is not judgment in the ordinary sense (Kant 1793, Religion. Ak. 6:186). He explains the difference quite explicitly:

*It is thus an instinct, and not merely a faculty of judgment.* Moreover, it is an instinct to direct and not to judge. The difference between a magistrate and one who judges is this: that the magistrate
judges *valide*, and actually put the judgment into effect according to the law; his judgment has the force of law, and is a sentence. A magistrate must not only judge, but also condemn or acquit. If conscience were an impulse to judge, it would be a cognitive faculty (Kant 1785, *Collins. Ak.* 27:352).

This explanation provides an answer to Moyar’s rightful question about the difference between first order judgment of deliberation and the second order judgment of conscience. Conscience is the instinct to ask the question “Do you believe that the action you are about to perform is the right (moral) action?” and it urges reason to scrutinize the first order judgment. Conscience sets the stage for the scrutinization. In this sense, it is similar to a court of appeals where the judgments of a lower court are reviewed diligently. “Thus understanding, judgment and reason are operative in the progress: conscience here reinforces awareness that the subject is in a situation that is governed by the laws of duty” (Kant 1793, *Vigilantius. Ak.* 27:617). Therefore, conscience is not the comprehensive judgment of action, but it is the drive to have comprehensive judgment of action. Since it summons us before the judgment seat, in regard to the lawfulness of our actions and “judge *valide*” the judgments of the reason, it seems to be the one who is performing the judgments. However, conscience is the cause of this second order judgment. The following text also supports this reading:

The duty of conscience presupposes, rather, that an action be legitimate or right; in this conception, conscience is regarded as a *potestas judicaria*, just as it is called *potestas legislatoria* and *executoria*, since it is really based on determining rectitude as such, on judging the *factum* by

---

60 In “Religion”, Kant provides a similar explanation: “Conscience does not pass judgment upon actions as cases that stand under the law, for this is what reason does so far as it is subjectively practical (whence the *casus conscientiae* and casuistry, as a kind of dialectic of conscience). Rather, here reason judges itself, whether it has actually undertaken, with all diligence, that examination of actions (whether they are right or wrong), and it calls upon the human being himself to witness for or against himself whether this has taken place or not” (Kant 1793. Ak. 6:185).
the laws of duty, and on establishing the *effectus a lege determinatorum et applicatorum*, and deliberately adopts conscience as a valid imputation of our actions. All this however, belongs to practical reason (Kant 1793, *Vigilantius. Ak*. 27:616).

Conscience has the authority to call for a scrutinization of our first order moral judgments because it has a judiciary power, which arises from the consciousness of the law of duty and being the subjective principle of being accountable to God.

This reading also supports Kant’s claim about the absurdity of an erroneous conscience, as we will see below. Since the second order judgment is undertaken by practical reason, the objective truth of the conclusion is not guaranteed. Indeed, this is not a requirement of conscience. Conscience is only responsible for calling a further investigation whether the subject is certain about the truthfulness of his action and “providing an awareness of having undertaking the examination with great thoroughness” (Kant 1793, *Vigilantius. Ak*. 27:619). Since one cannot be wrong about this call and awareness, conscience cannot err. As long as this investigation is conducted diligently, which amounts to acting in accordance with his conscience, “then as far as guilt or innocence is concerned nothing more can be required of him” (Kant 1797, *MM. Ak*. 6:401).

3.4 On the Possibility of an Erring Conscience

Erring conscience is a recurring theme in his writings on conscience. In his earlier treatments of erring conscience, there is a room for error about human law. In Herder and Collins lectures, Kant’s approach to the issue reminds us of medieval discussion of infallibility of *synderesis*. *Synderesis* is infallible because the knowledge of the natural law directs it to the good but *conscientia* may be mistaken because it draws particular
conclusions from the natural law and in the reasoning process, senses may deceive or reason may err. Kant provides a similar explanation:

The difference between the correct and the errant conscience lies in this, that error of conscience takes two forms, *error facti* and *error legis*. He who acts according to an errant conscience is acting conscientiously and if he does so, his action may be defective, but cannot be imputed to him as a crime. ...In regard to his natural obligations, nobody can be in error; for the natural laws cannot be unknown to anyone, in that they lie in reason for all...(Kant 1785, Collins. *Ak.*27:355).

It is not possible to be mistaken about the natural law. An erroneous conscience is possible only when there are mistakes in the legislation of positive laws and in the processing of the facts due to a “*conscientia erronea*”.

However, a few years later, in a journal article “On the Miscarriage of all Philosophical Trials in Theodicy” (“Theodicy”), Kant changes his mind about the issue and claims that “an erring conscience is an absurdity” (Kant 1791, “Theodicy”. *Ak.* 8:268). In the “Concluding Remark” section, while he is emphasizing the importance of sincerity in matters of faith, he also gives an account of conscience:

One cannot always stand by the truth if 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. For in the first instance we compare that 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)...We call this truthfulness “formal conscientiousness”; “material conscientiousness” consists in the caution of not venturing anything on the danger that might be wrong, whereas “formal conscientiousness” consists in the consciousness of having applied this caution in a given case (Kant 1791. *Ak.* 8: 267).

According to this account, conscience has a twofold function. First, *to make sure* that understanding did its job properly to find out the right thing to do. Second is *to check* whether the first task is fulfilled. In other words, conscience is supposed to bear witness
to the truthfulness of one’s judgment about the right thing to do. It is always possible to err in the judgment about the right thing to do. However, it is not possible to be mistaken for one’s own testimony whether she in fact believes what is to be right. Therefore, an erring conscience is an absurdity, according to Kant.

In the Vigilantius lectures, he takes up the issue again and argues that the division between an errant and a sound conscience (“inter conscientiam eroneam et rectam”) is false because “when the consciousness of what constitutes our duty is coupled with the judgment that a thing is right or wrong, though in itself it was impermissible or right, such an understanding merely judges erroneously” (Kant 1793, Vigilantius. Ak. 27:614). Kant is very clear that it is the duty of understanding to decide whether a thing is right or wrong and this decision may be fallible but one cannot be mistaken about her consciousness of her will, her disposition to do right.

Finally, in the MM, Kant summarizes his previous ideas almost with the same words. He claims that though there are “various divisions of conscience” he will deal only with erring conscience in the current section. This remark shows the importance he attributes to the issue. The reason why he takes up the issue again and again is that when he claims that “an erring conscience is an absurdity” he tries to reveal at least one of the ways conscience functions. It is not possible for conscience to err because one of the functions of conscience is to pass judgment on judgment as described below:

Conscience can also be defined as the moral faculty of judgment, passing judgment on itself…Conscience does not pass judgment on actions as cases that stand under the law, for this is what reason does so far as it is subjectively practical… Rather, here reason judges itself, whether it has actually undertaken, with all diligence, that examination of actions (whether they are right or
wrong), and it calls the human being to himself to witness for or against himself whether this has taken place or not (Kant 1793, *Religion. Ak.* 6:187).

Hence, primarily conscience is the *very act of submitting* our judgments for further investigation. Therefore, “while I can indeed be mistaken at times in my objective judgment as to whether something is a duty or not, I cannot be mistaken in my subjective judgment as to whether I have submitted it to my practical reason (here in its role as judge) for such a judgment” (Kant 1797, *MM. Ak.* 6:401). It is the duty of reason to judge whether an action is right or wrong on the basis of objective criterion and it is a duty of conscience to ask reason to scrutinize its first order judgment and it is not possible to be wrong whether one asks this or not.

Thomas Hill proposes an alternative interpretation for the impossibility of an erring conscience. He argues that:

> If on the one hand, we did scrutinize our act by our moral standards, we would have known this easily by introspection, and if so, conscience would have ‘involuntarily’ reached its verdict and (if appropriate) imposed its sentence. Mistakes here are apparently assumed to be impossible because what we compare is all ‘internal’: our conception of our act and our moral judgment regarding its rightness or wrongness. But if we did not submit our act to our moral standards, we did not make any prior moral judgment on the particular act, and so our conscience (which presupposes such judgments) never operated and so cannot have yielded a false verdict. Mistakes due to bad memory of our past acts and/or deliberations, misjudgments of objective duty, self-deceived conceptions of our acts, and the like are not counted as errors of conscience but as failures antecedent to its operation (Hill 2002, p.303 ff).

However, his interpretation is not explanatory enough. Even if “what we compare is all internal”, comparison is a process that may produce errors regardless of the things compared. So, working on “internal” content would not save conscience from error.
Richard Sorabji suggests that one way of saving conscience from error is to identify it with the supervisory role of the *praetor* in the Roman Republic, who is responsible for the proper conduct in the court of roles of other officials. He claims that “conscience is expected only to make sure that it has submitted (or, perhaps, supervised the submission of) conduct for examination and that examination (carried out by other, fallible entities) has been thorough (Sorabji 2014, p. 183). However, even Sorabji admits that when conscience supervises the judicial activities of the other fallible entities, it is carrying on their activities *itself* and this process may produce errors, as well. Hence, conscience is not immune from mistakes, even as a supervisor.

Allen Wood comes to a similar conclusion that I draw above. He claims that:

Since Kant does not identify conscience with moral judgment, he declines to infer from such cases that conscience can err. For Kant, conscience is rather the process of moral reflection that makes use of such moral judgments in delivering on myself a verdict of guilt or acquittal for some action I have done, or am contemplating. The duty of conscience is therefore the duty to engage in a kind of second-order reflection, judging that one has applied moral judgments properly to oneself … For conscience to err, therefore, *would be for me to be hold mistakenly that I have submitted myself and my action to this process when in fact I have not*. It is this error that Kant apparently regards as impossible. (Wood 2009, p. 14. Emphasis added).

The *very act of submitting* our judgments for further investigation is the primary function of conscience. It takes up this function because of its susceptibility to recognize duty. However, this is just one of the tasks conscience undertakes. The other one is to carry out the whole process of conducting a further investigation of one’s action. Kant describes this process with an internal court metaphor, as we have seen above.
3.5 Conclusion

Kant overcomes several problems medieval thinkers tried to solve regarding conscience, in his account of conscience. Kant does not adhere to the medieval distinction between synderesis and conscientia, which enables medieval thinkers to save a part of conscience, namely synderesis, from errors. Instead, Kant works on a single term, conscience (Gewissen). He passes some of the functions of synderesis and conscientia to the faculty of understanding and judgment. Conscientia draws particular conclusions from the general propositions and concludes what is the right thing to do. According to this account the source of moral knowledge is synderesis. However, the content of synderesis is ambiguous. Do we know all the general moral propositions? If so how general are they? Do we have innate knowledge of them or acquire them? No medieval account of conscience provides convincing answers to these questions.

According to Kant, the source of moral knowledge is the practical reason. Understanding provides rules, which determine objectively the rightness or wrongness of an act and then judgment brings cases under laws. Kant distinguishes these activities from the specific functions of conscience. His contribution to the understanding of conscience can be summarized as follows:

1. It is the “consciousness of what duty is” (Kant 1793, Vigilantius. Ak. 27:614).
   Conscience is a natural disposition (Kant calls it instinct, as well) to do right. This role is similar to synderesis. It is a natural disposition of reason, which always directs towards good.

2. Conscience brings every moral judgment to the court of appeals to be reviewed according to two criteria; first whether the first order judgment has been carried
out properly, and second whether the subject is certain about the decision. Conscience calls for a second order judgment and bearing witness to the process at the end pronounces the sentence. If the trial takes place before the subject undertakes an action, conscience warns him about the unlawfulness of his action. If it is a trial about the past deeds, conscience condemns or acquits the subject. *Synderesis* also “murmurs back in answer to sin” but no medieval thinker explains the mechanism how *synderesis* detects any transgression of the moral law. According to the medieval account, *conscientia* carries out the first order judgment. It is unclear how *synderesis* becomes aware of the failures of this judgment of *conscientia*.

3. An erring conscience is an absurdity according to Kant, because conscience is the very act of submitting our first order judgments for further investigation and it makes sure that one actually judges an action according to the law. It is not possible to be mistaken about these acts. *Synderesis* is also infallible but for a different reason. It is a natural disposition of reason, which always directs towards good and it is impossible for *synderesis* to go astray.

4. Finally, both medieval thinkers and Kant could agree with an important passage in St. Paul’s Epistle to the Romans (2:14–15). The gentiles do not have the written Jewish law, but they find an unwritten law in their hearts. For medieval thinkers,

---

61 It is also remarkable to see the similarity between the terms Paul and Kant uses, when we compare two text in German:
Rom 2.14 – 15 in the Luther Bible: “Denn so die Heiden, die das Gesetz nicht haben, doch von Natur tun des Gesetzes Werk, sind dieselben, die weil sie das Gesetz nicht haben, sich selbst ein Gesetz, als die da beweisen, des Gesetzes Werk sei geschrieben in ihren Herzen, sintemal ihr Gewissen ihnen zeugt, dazu auch die Gedanken, die sich untreineinander verklagen oder entschuldigen...”
Ak. 6: 437–8: “Das Bewußtsein eines inneren Gerichtshofes im Menschen (vor welchem sich seine Gedanken einander verklagen oder entschuldigen) ist das Gewissen.”
the law is put there by God. Kant thinks that our own reason gives us the law. Morality can be understood only if we see that each of us is equally a law giving member of the group of those, who must also obey the moral law. Hence each of us is responsible for both the legislation of the law and obeying it. In this way, Kant re-secularizes the concept of conscience, which originates as a secular one.
CHAPTER IV

Nietzsche

4.1 Introduction

As shown in the previous chapter, according to Kant, only the feelings that result solely from the activity of pure reason, including remorse and bad conscience are non-pathological. However, Nietzsche thinks that it is precisely these Kantian “non-pathological” feelings that are truly pathological. Nietzsche gives a “historical” account of these feelings in the second essay of On the Genealogy of Morals (GM), “‘Guilt,’ ‘Bad Conscience,’ and the Like”. However, if we examine Nietzsche’s account of conscience closely we could find some similarities between Nietzsche and Kant. The apparent controversy arises from Nietzsche’s emphasis on bad conscience throughout the essay. He devotes only one out of twenty-five sections to conscience. The remaining sections are about the historical development of bad conscience, which is considered by Nietzsche as an early stage of conscience. The aim of this chapter is to discuss whether Nietzsche’s genealogy of bad conscience really challenges the traditional account of conscience and to show the similarities between Kant’s and Nietzsche’s account of conscience by providing an interpretation of the second essay of the Genealogy.

Unlike the first essay “‘Good and Evil,’ ‘Good and Bad’”, the second essay has not attracted as much scholarly attention. This is due to the exegetical difficulties of the essay. Aaron Ridley even argues it is almost impossible to bring together Nietzsche’s

---

62 Important exceptions are Clark 1994; Ridley 1998; Risse 2001. Although all of them provide insightful comments on some parts of the second essay, none of them addresses the issues of my chapter.
various ideas on conscience harmoniously because “the use that he makes of that concept is far from transparent. The distinction and the function of the distinction, between good and bad conscience is perplexing; and the relation between bad conscience and ressentiment is scarcely less so” (Ridley 1996, p.1). Nietzsche takes precautions against this type of criticism in the last section of the preface. He warns that “if this book is incomprehensible to anyone and jars on his ears, the fault, it seems to me, is not necessarily mine. It is clear enough, assuming, as I do assume, that one has first read my earlier writings…” (GM, Preface, 8). As a matter of fact, just after the title page, he writes: “A sequel to My Last Book, Beyond Good and Evil, Which It Is Meant to Supplement and Clarify”. Hence any attempt to interpret “‘Guilt,’ ‘Bad Conscience,’ and the Like,” has to take Nietzsche’s previous works and his understanding of morality into consideration in order to make sense of the text.

4.2 Nietzsche and Morality

Nietzsche is considered to be an immoralist, even he claims to be so (Ecce Homo IV, 2-4; The Birth of Tragedy, P, 5). This qualification needs to be clarified; exactly in what sense is he an immoralist? Many commentators argue that Nietzsche rejects a particular kind of morality or a particular theory or conception of morality. However, Maudemarie Clark shows that “they rely on their own understanding of the concept of morality to make sense of Nietzsche’s immoralism” (Clark 1994, p. 19). For example, since Foot thinks

---

63 Unless otherwise stated I use Walter Kaufmann’s translation and cite the text by referring to the chapter and section numbers. Basic Writings of Nietzsche. Translated and edited by Walter Kaufmann. The Modern Library: New York, 1992.

64 Among others, most prominent defenders of this interpretation are Philippa Foot, Alexander Nehamas, Frithjof Bergmann, Walter Kaufmann, Arthur Danto, Robert Salomon, John Wilcox, Richard Schacht and Frederick Olafson.
that justice is essential to morality she concludes that Nietzsche is an immoralist because he rejects justice. Nehamas and Bergmann do the same thing with reference to their understanding of the universalizing one’s values and freedom. Clark proposes a new criterion to evaluate Nietzsche’s immoralism: Nietzsche’s own analysis of the concept of morality. Although it is not so different from Clark’s interpretation, I attempt to give my own explanation of Nietzsche’s understanding of morality briefly, so that it will be clearer what exactly he is rejecting. In addition, an understanding of Nietzsche’s view of morality enables us to make better sense of the second essay.

Nietzsche’s understanding of historical method can provide a clue for his analysis of the concept of morality. In the middle section of the second essay, he proposes a principle for “historiography of any kind”:

The cause of the origin of a thing and its eventual utility, its actual employment and place in a system of purposes, lie worlds apart; whatever exists, having somehow come into being, is again and again reinterpreted to new ends, taken over, transformed and redirected by some power superior to it; all events in the organic world are subduing, a becoming master, and all subduing and becoming master involves a fresh interpretation, an adaptation through which any previous “meaning” and “purpose” are necessarily obscured or even obliterated. However well one has understood the utility of any physiological organ (or of a legal institution, a social custom, a political usage, a form of art or in a religious cult), this means nothing regarding its origin: however uncomfortable and disagreeable this may sound to older ears – for one had always believed that to understand the demonstrable purpose, the utility of a thing, a form, or an institution, was also to understand the reason why it originated – the eye being made for seeing and hand being made for grasping (GM II, 12).

Obviously, Nietzsche adapts a common sense understanding of Darwinism65 to history and criticizes Lamarckian understanding of history, which assumes that any change in the

---

organism is purposeful. However, Nietzsche replaces Darwin’s principle of evolution, namely “survival of the fittest” with “will to power” as the driving force of life. On the basis of this modified version of Darwinism he gives an account of morality by examining some practices and meanings and purposes attributed to them in the course of time. By doing so, he tries to show that morality has not always been with human beings, rather it is an emergent “synthesis” (GM II,13) that ties different procedures together that were successful enough to transform an animal into the one “with the right to make promises” (GM II, 1), in other words into a moral being. As I will explain below, he is against the price we pay for becoming moral agents and claims that it is possible to get rid of it. In this sense, he can be considered as an immoralist.

4.3 Development of the Bad Conscience

The structure of *Genealogy*, which consists of three independent essays, supports the above reading. Nietzsche consciously chooses this structure as he explains on a postcard to Franz Overbeck from January 4, 1888:

> For the sake of clarity, it was necessary artificially to isolate the different roots of that complex structure that is called morality. Each of these three treatises expresses a single *primum mobile*; a fourth and fifth are missing, as is even the most essential (‘the herd instinct’) – for the time being, the latter had to be ignored, as too comprehensive, and the same holds for the ultimate summation of all those different elements and thus a final account of morality.66

The three treatises discuss different strands in the development of morality. These separate discussions are done in abstraction from their actual interaction. The same is true for the structure of the second essay. It seems to consist of various ideas on conscience, which cannot be brought together consistently, as Ridley criticizes above. However,

---

66 Originally in Mathias Risse 2001, p.55
Nietzsche aims to show that not only morality but also its components – like any other concept – in our case, bad conscience is a synthesis of different procedures and meanings. Some of these procedures and meanings follow each other historically, others take place at the same time. For this reason, to bring unity to these strands is a difficult task, although not impossible, especially if we keep in our mind Nietzsche’s remark on the issue:

For it must be obvious which color is a hundred times more vital for a genealogist of morals than blue: namely gray, that is, what is documented, what can actually be confirmed and has actually existed, in short the entire long hieroglyphic record, so hard to decipher, of the moral past of mankind! (GM, P 7)

Given all this complexity, what I offer below is an attempt to explain Nietzsche’s account of bad conscience and hence conscience. I then show his contribution to the traditional approach by a challenging of its presuppositions about the content and function of conscience.

It may be surprising to see that Nietzsche’s treatment of bad conscience starts with a remark on making promises. Indeed, this is quite meaningful, for being accountable is one of the prerequisites of being a moral agent and keeping promises is a definite sign of being accountable. He claims that “to breed an animal with the right to make promises” (GM II,1) is a real problem regarding man because accomplishing this task would require transforming human nature into its opposite. Hence, the first section of this treatment of bad conscience is the introduction of the essay, where Nietzsche informs us about his project: to explain the processes through which human beings become moral agents. In the second section, he describes the outcome of “this tremendous process”: the sovereign individual. Starting with the third section, he identifies and explains the
components of this process. Almost each section concentrates on the development of one of the components or strands, if we follow Clark’s terminology: “Nietzsche suggests that concepts influenced by history are like ropes held together by the intertwining of strands, rather than by a single strand running through the whole thing” (Clark 1994, p.22).

Actually, each essay tries to disentangle the various strands, which may have become so tightly woven together, with a different point of focus. The focus of the second essay is bad conscience and guilt as the title indicates, specifically, their emergence and their role in the constitution of the moral agent.

Nietzsche identifies acquiring a memory as the first step in the process of acquiring the ability to make promises and hence having a bad conscience. According to him, the human animal by nature is a forgetful one. However, once they realized the effects of inflicting pain on remembering they developed most fearful mnemotechnics to make sure that one remembers “five or six ‘I will not’s,’ in regard to which one had given one’s promise to participate in the advantages of society” (GM II,3).

Ridley quickly concludes that creating a memory has to be in the form of imposing customs and asks “what capacities must someone have if he is to impose a custom on others[?]” (Ridley 1996, p.3). The answer for Ridley is obvious: power. However, Ridley claims that this is not enough. He claims that “the imposer of the custom must himself have a memory of the will and have become calculable which means in turn that he must have been subjected to custom and punishment” (Ridley 1996, p.3). According to Ridley, the problem is not the circularity of the account but the inconsistency of Nietzsche’s claims about the imposers of the customs: the nobles who are depicted as “pack of blond beasts of prey”, whose “work is an instinctive creation and
imposition of forms; they are the most involuntary, unconscious artists there are” (GM II, 17). Therefore, either they cannot have a memory of the will and become calculable and hence they are not capable to impose any custom on others, or these beasts were imposed upon by customs and hence cannot have the qualities Nietzsche attributed to them.

There is a problem with Ridley’s objection. Inflicting pain on others and then realizing its effects does not have to be in the form of custom. According to Nietzsche, human beings are hostile, cruel and they enjoy persecuting, attacking and destroying by their very nature (GM II, 16). Therefore, Nietzsche is consistent when he claims that:

Some pack of blond beasts of prey…come like fate, without reason, consideration, or pretext; they appear as lightning appears, too terrible, too sudden, too convincing, too “different” even to be hated. Their work is an instinctive creation and imposition of forms; they are the most involuntary, unconscious artists there are – wherever they appear something new soon arises, a ruling structure that lives, in which parts and functions are delimited and coordinated, in which nothing whatever finds a place that has not first been assigned a “meaning” in relation to the whole. They do not know what guilt, responsibility, or consideration are… It is not in them that the ‘bad conscience’ developed, that goes without saying but it would not [have?] developed without them (GM II, 17).

So, when these beasts realized that by inflicting pain on others (in accordance with their nature) they were able to change them in a way convenient for themselves, they created the memory unconsciously. This realization “comes into being somehow” and “has not first been assigned a ‘meaning’ in relation to the whole” but it “is again and again reinterpreted to new ends, taken over, transformed” (GM II, 12). In the course of time, more elaborated mnemotechnics emerged to shape and control human being’s instincts. Nietzsche’s evidence is the process of breeding Germans as a “nation of thinkers”. He claims that by using harsh methods, such as stoning, breaking on the wheel, piercing with stakes, tearing apart or trampling by horses, Germans learned to “master their mob
instincts and its brutal coarseness” and they become the nation in Europe that has “the maximum of trust, seriousness, lack of taste and matter-of-factness” (GM II, 2).

Once Nietzsche has explained how by nature a forgetful animal is transformed into the one who remembers what to do and not to do in order “to participate in the advantages of society,” he is ready for an explanation of another strand: transformation of the material concept Schulden (debts) into the moral concept Schuld (guilt).

The concept debt emerges out of the ordinary exchanges among human beings such as buying, selling, barter, trade, and traffic. In return for goods or services something has to be given to the provider. Any kind of exchange is based on this principle of fairness. Nietzsche does not give an account for this principle, he takes it for granted. If one party is unable to pay for something he gets at the moment of the exchange then the other party wants the debtor to make a promise to repay his debt in a given time. So, “the debtor makes a contract with the creditor and pledged that if he should fail to repay he would substitute something else he ‘possessed,’ something he had control over; for example, his body, his wife, his freedom, or even his life” (GM II, 5).

Nietzsche points out a strange implementation of this principle: inflicting pain and indignity upon the body of the debtor, by for example, cutting from it as much as seemed commensurate with the size of the debt. Nietzsche observes that making the debtor suffer balances debts to the extent that “to make suffer was in the highest degree pleasurable” (GM II, 6). So that “the injured party exchanges for the loss he had sustained, including

---

67 This claim sounds Hobbesian but according to Nietzsche there is not and cannot be a conscious act of entering a contract: “I think that sentimentalism which would have it begin with a ‘contract’ has been disposed of. He who can command, he who is by nature ‘master,’ he who is violent in act and bearing – what has he to do with contracts!” (GM II, 17). What he reveals here is that the rationalization of human beings who are forced to act in certain manner.
the displeasure caused by the loss, an extraordinary counterbalancing pleasure: that of making suffer” (GM II, 6). He even claims that punishing someone for his debts by making him suffer has the nature of a festival. He supports this claim with historical examples and points out that throughout the history celebrations of any kind are accompanied by deeds of cruelty: “Without cruelty there is no festival” (GM II, 6).

Nietzsche elaborates on human beings’ relation to cruelty in his other writings as well and claims that cruelty is in the essence of human being:

What constitutes the painful voluptuousness of tragedy is cruelty; what seems agreeable in so-called tragic pity, and at bottom in everything sublime, up to the highest and most delicate shudders of metaphysics, receives its sweetness solely from the admixture of cruelty. What the Roman in the arena, the Christian in the ecstasies of the cross, the Spaniard at an auto-da-fe or bullfight, the Japanese of today when he flocks to tragedies, the laborer in a Parisian suburb who feels a nostalgia for bloody revolutions, the Wagnerienne who ‘submits to’ Tristan and Isolde, her will suspended – what all of them enjoy and seek to drink in with mysterious ardor are the spicy potions of the great Circe, ‘cruelty’ (BGE 229).

On the basis of these ‘facts” he derives “the principle to which even the apes might subscribe” that “to see others suffer does one good, to make others suffer even more” (GM II,6). Given this principle, when the creditor makes the debtor suffer, he takes pleasure, which compensates his loss.

According to Nietzsche, the relationship between creditor and debtor is the basic form of personal interaction and the principle, which governs this relationship is the basic principle of justice that is “everything has its price and; all things can be paid for”. The same is true for the relationship between community and its members. By being a member of a community one enjoys the advantages of a community such as being protected, cared for, living in peace and trustfulness, without fear of certain injuries and
hostile acts. In return for all these goods and benefits, one has to pledge one’s loyalty to the community. When this pledge is broken, the community, as the creditor will ask for a repayment. This repayment is not only for the direct harm but also for the broken contract since the lawbreaker attacks the very foundation of a community: law and order. The price for the broken law is being outside of the law. The community throws the debtor back into the “savage and outlaw state against which he has hitherto been protected” (GM II, 9). Nietzsche states that at this level of civilization, this “punishment” is a copy of the normal attitude toward a “hated, disarmed, prostrated enemy”.

Nietzsche puts punishment in quotation mark because he thinks that the notion of punishment needs further analysis if he is to make sense of bad conscience. He therefore applies the historical method, which is explained above briefly. He claims that punishment can be analyzed with respect to its two aspects: a strict sequence of procedures that is the enduring aspect, and the meaning, the purpose, the expectation associated with the performance of such procedures, namely the fluid aspect. The procedures exist before their employment in punishment, they come into being somehow, maybe as an expression of human beings’ cruel instincts. The meaning of punishment is projected and interpreted into the procedure. So, contrary to common understanding the procedures are not invented for the purpose of punishing. Furthermore, the concept of punishment does not have one meaning but it has a “whole synthesis of meanings”, such as:

- punishment as a way of rendering someone harmless, as a prevention from further harm;
- punishment as compensation for the damage to the person injured, in some form or other (also in the form of emotional compensation);
- punishment as isolation of some upset to an even balance in order to avert a wider outbreak of the disturbance;
- punishment as a way of bringing fear to those
who determine and carry out punishment; punishment as a sort of compensation for the advantages which the law breaker has enjoyed up until that time (for example, when he is made useful as a slave working the mines); punishment as a cutting out of a degenerate element (in some circumstances an entire branch, as in Chinese law, and thus a means to keep the race pure or to sustain a social type); punishment as festival, that is, as the violation and humiliation of some enemy one has finally thrown down; punishment as a way of making a memory, whether for the man who suffers the punishment—so-called “reform”—or whether for those who witness the punishment being carried out; punishment as the payment of an honorarium, set as a condition by those in power, which protects the wrong doer from the excesses of revenge; punishment as a compromise with the natural condition of revenge, insofar as the latter is still upheld and assumed as a privilege by powerful families; punishment as a declaration of war and a war measure against an enemy to peace, law, order, and authority, which people fight with the very measures war makes available, as something dangerous to the community, like a contract breaker with respect to its conditions, like a rebel, traitor, and breaker of the peace (GM II, 13).

With this overdetermined, incomplete list, Nietzsche wants to show how the same procedure can be employed, interpreted, adapted to ends that differ fundamentally and hence “the meaning” of punishment is “uncertain, supplemental and accidental”.

To awaken the guilty feeling in the guilty person and thus the development of conscience is thought to be among the purposes of punishment. Nietzsche rejects this assumption. He claims that punishment had a contrary effect. It hinders the development of the feeling of guilt because first through punishment the guilty person compensates for his crime and second the judicial and executive procedures are not so different than the crimes they punish, such as spying, deception, bribery, violence, robbery, torture, imprisonment, murder, etc. Nietzsche concludes that neither judges or punishers nor the person who gets the punishment have a conception of guilt. What they see is “an instigator of harm, with an irresponsible fate” (GM II, 14). And the “mischief-makers overtaken by punishment have for thousands of years felt in respect of their “transgressions” just as Spinoza did: ‘here something has unexpectedly gone wrong,’ not:
‘I ought not to have done that.’ They submitted to punishment as one submits to an illness or to a misfortune or to death” (GM II, 15). To make sense of this observation, Nietzsche cites Spinoza: “A man is as much affected pleasurably or painfully by the image of a thing past or future as by the image of a thing present” (Ethics III, propos XVIII, schol. I-II). For Nietzsche, the question is: which memories or expectations change human beings’ evaluation of their transgressions, so that they replace “here something has unexpectedly gone wrong,” with “I ought not to have done that?” In other words, what makes human beings develop a “bad conscience?”

Nietzsche provides a “provisional statement” of his hypothesis concerning the origin of the “bad conscience”. He considers bad conscience as the serious sickness; a reaction to the transition from wilderness to civilization. He compares this process to the transformation of sea animals into the land animals, which are:

well adapted to the wilderness, to war, to prowling, to adventure: suddenly all their instincts were disvalued and suspended…They felt unable to cope with the simplest undertakings; in this new world they no longer possessed their former guides, their regulating, unconscious and infallible drives: they were reduced to thinking, inferring, reckoning, coordinating cause and effect; they were reduced to their “consciousness”, their weakest and most fallible organ! (GM II, 16).

According to Nietzsche, by nature human beings are hostile, cruel, take joy in persecuting, in attacking, and in destruction. The basis of this claim is Nietzsche’s commitment to the principle that life is primarily will to power. He explicitly states this principle in Beyond Good and Evil: “Life itself a process of appropriating, injuring, overpowering the alien and the weaker, oppressing, being harsh, imposing your own form, incorporating, and at least, the very least, exploiting” (BGE, 259). However, civilization forces human beings to control these instincts to protect itself. As a result,
“all instincts that do not discharge themselves outwardly turn inward” (GM II, 16).

Nietzsche calls this process internalization and claims that in the beginning the entire inner world was very thin but the more outward discharge was inhibited the more it was expended and extended and became the “soul” of human being.

Nietzsche gives a historical account for this transition. He claims that the transition was not a “gradual or voluntary one” but “a break, a leap, a compulsion” and it was instituted and carried out by acts of violence” (GM II, 17). According to Nietzsche, the oldest state, which gives rise to this transition, was formed by the attack of “some pack of blond beasts, a conqueror and master race, which organized for a war and with an ability to organize” on formless and nomad people. As explained above, the conquerors used violent techniques to shape these nomads in a way that is convenient for themselves. As a result, even though the conquerors did not know what guilt, responsibility, or consideration are, they made others to develop them and push back their instincts. The instincts that are pushed back, repressed are discharged on the subject itself. This is the origin of bad conscience according to Nietzsche. Although the conquerors did not have it, “it would not have developed without them” (GM II, 17).

---

68 For a discussion about “blond beasts, master race” see Kaufmann’ Nietzsche, Ch. 10 “The Master Race.” Kaufmann claims that master race refers to the Homeric Greeks. If this is the case then Nietzsche’s account is not historically accurate because the oldest states in the history were founded in Mesopotamia. I think Nietzsche depicts the nature of the founders of the first states rather than their race as he explicitly states in BGE 257: “Let us be not deceived about how every higher culture on earth has begun! Men whose nature was still natural, barbarians in every terrible sense of the word, predatory people who still possessed an unbroken will and lust for power, threw themselves on weaker, more civilized, more peaceful races of tradesmen perhaps, or cattle breeders; or on old mellow cultures in which the very last life-force flaring up in brilliant fireworks of spirit and corruption. The noble cast always started out as the barbarian caste. Their supremacy was in psychic, not physical strength – they were more complete people (which at any level amounts to saying ‘more complete beasts’).
Towards the end of the essay, Nietzsche starts to bring the different strands that are woven into the concept of bad conscience. The first one is the “interpretation” of debtor–creditor relationship into the relationship between the present generation and its ancestors. He states that starting with the primeval times, the living generations always recognized a juridical duty toward earlier generations, especially toward the founders of the tribe. For they believe that only through the sacrifices and accomplishments of the ancestors the tribe exists. Hence they owe their existence to their ancestors, a debt that can hardly be paid off. Even if they try to repay their debt with sacrifices, feasts, honors and above all with obedience, the suspicion about the adequacy of the repayment remains, in fact, increases. The more powerful and victorious a tribe becomes, the consciousness of indebtedness toward its ancestors increases. With the decline of a tribe its ancestors are not be seen as a source of fear and debt. What follows from this kind of logic is that with the help of imagination is triggered the fear of the ancestor of the most powerful tribes, transfiguration of the ancestor into a god. Nietzsche claims that this could be the origin of the gods.

Nietzsche gives another account of the origin of gods in previous sections, when he deals with the internalization of cruelty. He claims that once human beings learned to be ashamed of their cruelty they turned it inwards. They suffered this instinctive cruelty but they needed a reason to endure it because what makes suffering unbearable is the senselessness of it. Nietzsche concludes that “to abolish hidden, undetected, unwitnessed suffering from the world” human beings were compelled to invent the gods (GM II, 7).

These two different accounts about the origins of gods seem to be inconsistent but actually they refer to the different aspects of the concept of god: in the first one god is the
creator and the owner of the universe to whom we owe everything we have, including our bodies and souls; in the second, god is omniscient who knows everything, including our deepest secrets. The idea of a god or an inner judge to whom we have to render an account for our deeds is an important part of the conceptualization of conscience. Nietzsche likewise emphasizes that our feeling of indebtedness/ guiltiness to the god plays a crucial role in the understanding of conscience. For this reason, he dwells further on the history of the development of this feeling.

Nietzsche states that the guilty feeling of indebtedness to the divinity continued to grow for several millennia and reached its peak with the advent of the Christian God. “[A]s the maximum god attained so far, [it] was therefore accompanied by the maximum feeling of guilty indebtedness on earth” (GM II, 20).

For these reasons, we see that Simon May’s critique of Nietzsche does not hold. May identifies problems with Nietzsche’s equation of debt and guilt:

1. Debt denotes an obligation, which one may or may not be capable of discharging and which one may or may not recognize; whereas guilt denotes a feeling consequent upon failure to discharge an obligation that one does recognize. Thus, not all debt engenders guilt: specifically, not debt that one can and does repay on time, or that one never recognized as repayable.

2. Nietzsche cannot be right to regard the feeling of guilt as simply a feeling of “personal obligation” (GM II, 8). Guilt is occasioned only by failing to honor what we take to be an obligation. It is not always necessary to experience it as impotence (May 1999, p. 59).

May misses a very crucial point Nietzsche makes about the transformation of debt into guilt: the moralization of the term Schuld. Following the advent of the Christian God, human beings recognized the fact that it is not humanly possible to repay their debt to this

---

69 The legal term Shuld/debt became the moral term Schuld/guilt.
maximum god. Therefore, they admitted their failure and impotence and hence developed the feeling of guilt. The bad conscience is the consciousness of this guilt.

Apparently, this is just the surface of the “whole truth” according to Nietzsche. He summarizes his findings towards the end of the essay:

You will already have guessed what has really gone on with all this and behind all this: that will to torment oneself, that suppressed cruelty of animal man who has been frightened back into himself and given an inner life, incarcerated in the ‘state’ to be tamed, and has invented bad conscience so that he can hurt himself, after the more natural outlet of this wish to hurt had been blocked, – this man of bad conscience has seized on religious presupposition in order to provide his self-torture with its most horrific hardness and sharpness. Debt towards God: this thought becomes an instrument of torture. In ‘God’ he seizes upon the ultimate antithesis he can find to his real and irredeemable animal instincts, he reinterprets these self-same animal instincts as debt/guilt before God (as animosity, insurrection, rebellion against the ‘master’, the ‘father’, the primeval ancestor and beginning of the world), he pitches himself into the contradiction of ‘God’ and ‘Devil’, he emits every ‘no’ which he says to himself, nature, naturalness and the reality of his being as a ‘yes’, as existing, living, real, as God, as the holiness of God, as God-the-Judge, as God-the-Hangman, as the beyond, as eternity, as torture without end, as hell, as immeasurable punishment and guilt (GM II, 22).

Commentators focus on this passage and try to interpret it without paying much attention to the previous sections, where Nietzsche lays down the strands that he weaves into the concept of bad conscience, and they develop their interpretation on the basis of what is on the surface. Thus, Clark argues that what transforms debt into guilt is “the ascetic ideal and its attendant ascetic conception of virtue”, which Nietzsche discusses at length in the third essay of the Genealogy (Clark 1994, p.30). Risse replaces the ascetic ideal with the Christian story and argues that “its acceptance gives rise to an entirely new sentiment, namely guilt, which is so strong that by itself it gives rise to a new kind of moral psychology (Risse 2001, p. 64).
However, it is clear on textual grounds that Nietzsche thinks that bad conscience emerged as a consequence of the internalization of cruel instincts. At first, it was the inner world, which was invented by human beings to exercise their cruel instincts on themselves, since they were repressed by the society. The idea of having debts towards God enables human beings to make sense of their self-torture: *I deserve this punishment because I am in debt for what I am, a complete failure with respect to what one is first and foremost, namely, God's creature. I am the ultimate antithesis of God because of my ineluctable animal instincts. I am guilty and reprehensible to a degree that can never be atoned for.* Hence, the indebtedness has turned into guilt and bad conscience is the consciousness of this unatonable guilt.

To sum up, the source of the bad conscience is the internalization of instincts; the content is the guilty feelings arises from the failure to respect the timeless, absolute, impartial standards, which are represented by the idea of God, and the function of the bad conscience is to make sense of the suffering.

Whenever, human beings fail to meet these standards, they inflict pain on themselves, so that they find an outlet to exercise their cruel instincts. The side effect of this process is the emergence of the moral agent. The inward directed cruelty of human beings is not a threat for the society anymore, the taming of the animal man is accomplished. “With the aid of the morality of mores and the social straitjacket, man was actually made calculable” (GM II, 2).

---

70 In the First Essay of the Genealogy Nietzsche gives an etymological account of the origins of moral agent by examining the term “good”. This is one of the strands, which is woven into the concept of morality and the moral agent. In the Second Essay he provides another account of the origins of moral agent by examining the concept of bad conscience. Similarly, in the Third Essay, he takes up the issue with reference to the ascetic ideal.
4.4 Emergence of Conscience

According to Nietzsche, the bad conscience is a sickness because it maintains human beings disloyalty to their true nature. Nevertheless, “it is a sickness as pregnancy is a sickness”, because the “prospect of an animal soul turning against itself“is an event and a spectacle too interesting to be played senselessly unobserved on some ridiculous planet.” Furthermore, the bad conscience contributes to the appearance of an animal on earth that “arouses interest, tension, hope”, as if through it “something . . . were being prepared, as though man were not an end but just a path, an episode, a bridge, a great promise” (GM II, 16).

Nietzsche observes that although it represents a painful and ugly growth, the bad conscience is not simply to be looked upon in disparaging terms; indeed, he speaks of the “active bad conscience”. It can be regarded as the “true womb of ideal and imaginative events”; through it an abundance of “disconcerting beauty and affirmation” has been brought to light (GM II, 18). Hence, it may give birth to something life enhancing: the sovereign individual: “the ripest fruit on its tree, like only to itself, having freed itself from the morality of custom, an autonomous, supra-moral (übersittlich) individual (because ‘autonomous’ and ‘moral’ are mutually exclusive)” (GM II, 2).

Nietzsche presumes that after the attainment of the maximum god, now human beings have gradually entered upon the reverse course: “the irresistible decline of faith in the Christian God” will be followed by “a considerable decline in mankind’s feeling of guilt” (GM II, 20). He even anticipates that “the complete and definitive victory of atheism might free mankind of this whole feeling of guilty indebtedness toward its origin” (GM II, 20). With the death of god mankind would be able to free itself from the
morality of custom and to go back to an innocent state, where there will be no burden of the “original sin.”

At the end of this process, when mankind frees itself from the morality of mores and unatonable guilt before god, it would be possible to attain the autonomous, supra-moral sovereign individual. The sovereign individual would make peace with his cruel instincts as well because he would affirm himself as he is and would not need to turn them inward anymore.

Nietzsche’s claim about the mutual exclusiveness of autonomy and morality sounds controversial and anti-Kantian, at first. However, when he describes the sovereign individual in the following lines, his sovereign individual is not so different from Kant’s moral agent:

This man who is now free, who actually has the right to make promise, this master of the free will, this sovereign… how could he, with his self-mastery, not realize that he has necessarily been given mastery over circumstances, over nature and over all creatures with a less enduring and reliable will? The ‘free’ man, the possessor of an enduring, unbreakable will, thus has his own standard of value… that is everyone who promises like a sovereign, ponderously, seldom, slowly, and is sparing with his trust, who confers an honor when he places his trust, who gives his word as something that can be relied on, because he is strong enough to remain upright in the face of mishap or even ‘in the face of fate’… (GM II, 2, emphasis added).

Both Kant and Nietzsche emphasize freedom from circumstances, nature and one’s own society and autonomy, which is being the author of one’s own rules. Keeping promises, i.e., being accountable and responsible is a must for both of them. The main difference, although not an unimportant one, is Nietzsche’s rejection of the universal law, on the basis of his rejection of reason as the source of action. For him, the source of action is primarily human beings’ instincts. However, the apparent controversy is not due to this rejection but their definition of morality. For Kant morality has nothing to do with
morality of customs. An action is moral if and only if it is done from duty, i.e. if it follows the categorical imperative. Whereas, for Nietzsche morality refers basically to the customs/mores that aims to constraint individual’s instincts. For this reason, Nietzsche claims that “autonomous’ and ‘moral’ are mutually exclusive.” An individual who is bound by customs cannot rule himself and Kant would completely agree with him on this issue.

Despite their different views on the human nature and hence morality, still Kant and Nietzsche share a similar view on conscience:

The proud knowledge of the extraordinary privilege of responsibility, the consciousness of this rare freedom and power over himself and his destiny, has penetrated him to his lowest depths and become an instinct, his dominant instinct: – what will he call his dominant instinct, assuming that he needs a word for it? No doubt about the answer: this sovereign human being calls it his conscience . . . (GM II, 2, emphasis added).

Although they disagree on what duty/responsibility is, conscience is the consciousness of duty/responsibility for both them. Remarkably both Nietzsche and Kant call it an instinct. Obviously, unlike Kant, according to Nietzsche, conscience has only a legislative power and works only prospectively. Judiciary and executive functions of conscience are completely missing in this account.  

71 Moral in the sense of mores.
72 In GM II, 6, Nietzsche writes: “In this area, that is, in the laws of obligation, the world of moral concepts "guilt," "conscience," and "sanctity of obligations" was conceived. Its beginnings, just like the beginnings of everything great on earth, were watered thoroughly and for a long time with blood. And can we not add that this world deep down has never again been completely free of a certain smell of blood and torture—(not even with old Kant whose categorical imperative stinks of cruelty . . . ).”. It is no doubt that according to Nietzsche, the obligation that is imposed by the categorical imperative is problematic because it is a dictate of reason. And reason is one of the detrimental outcomes of civilization: “Ah, reason, seriousness, mastery over emotions, the whole gloomy business called reflection, all these privileges and ceremonies of human beings—how expensive they were! How much blood and horror is the basis for all "good things." (GM II, 3). Nevertheless, Both Nietzsche and Kant are in line when they claim that conscience is the recognition of duty, even though they disagree what duty is.
4.5 Conclusion

Nietzsche’s detailed analysis of the bad conscience and brief description of sovereign conscience display different aspects of conscience that have been discussed in the previous chapters. In addition, he gives an account about the origin of conscience. No one before him attempted to give an account about it because they have taken it for granted, as a built in capacity of human beings. His genealogy of conscience provides an insight about the function of it, regardless of its being real or fictional.73

First, he shows that bad conscience as a feeling of guilt exists only if the agent has internalized some ethical standards and develops a sense of personal accountability to them. Bad conscience therefore functions as an internal alert which informs about a transgression and develops guilty feelings. So far, Nietzsche’s explanations are in line with the traditional accounts. However, the debtor-creditor relationship (GM II, 6) provides a fruitful model for the internalization of ethical standards – which is missing in the traditional accounts – so that it explains how a person feels obligated to follow them, even in the absence of external threats, rewards and monitoring.

Second, although Nietzsche admits that bad conscience and the processes that give rise to it differentiates human beings from the rest of animals, he regards these processes as bad. However, for the same reason, traditional accounts view conscience as

---

the most sublime part of human beings. Nietzsche believes that it is possible to have a sovereign conscience by getting rid of the moral standards, which are represented by the idea of god and customs. But how is it possible to have conscience without a “regulative idea”, with reference to which one can check the rightness of one’s deeds? He insists that sovereign individuals would not fail to live up to their values because by being the originators of these values, they are fully aware that they bound themselves. So, there are two consequences of this claim: first, Nietzsche makes the content of conscience arbitrary and second, it is not possible for a sovereign individual to judge himself for his misdeeds. The first one is widely accepted in today’s secular world but if we accept the second one then we need to change the meaning of conscience.

Since we are not sovereign individuals in the Nietzschean sense, who are likely to fail to live up their own standards and hence suffer from these failures, what are our options regarding conscience in a secular world? In the conclusion, I attempt to give an answer to this question.

74 At the end of the Second Essay, Nietzsche claims that the sovereign individual is an ideal which would be realized with the help of anti-Christ: “This man of the future, who will release us from that earlier ideal and, in so doing, from those things which had to grow from it, from the great loathing, from the will to nothingness, from nihilism—that stroke of noon and of the great decision which makes the will free once again, who gives back to the earth its purpose and to human beings their hope, this anti-Christ and Anti-nihilist, this conqueror of God and of nothingness—at some point he must come . . .” (GM II, 23). In the Third Essay of the Genealogy, Nietzsche gives a clue to interpret this passage: “All great things destroy themselves by an act of self-cancellation. That’s what the law of life wills, that law of the necessary "self-overcoming" in the essence of life—eventually the call always goes out to the law-maker himself, "patere legem, quam ipse tulisti" [submit to the law which you yourself have established]. That’s the way Christianity was destroyed as dogma by its own morality, that’s the way Christendom as morality must now be destroyed. We stand on the threshold of this event. After Christian truthfulness has come to a series of conclusions, it will draw its strongest conclusion, its conclusion against itself. . . Because this will to truth from now on is growing conscious of itself, morality undoubtedly dies. That great spectacle in one hundred acts, which remains reserved for the next two centuries in Europe, that most fearful, most questionable, and perhaps also most hopeful of all spectacles . . .” (GM III, 27) It can be concluded that transition from Christian morality to sovereign “ethics”. This is a process which enables human beings to re-evaluate their values and come up with the ones that are in line with their true nature and free from the burden of any guilt. (I would like to thank Dr. Jennifer Bates for pointing out this claim.)
Conclusion

We most typically experience conscience as a voice that tells us what we have done or what we are about to do is wrong and when we hear this voice we feel remorse or apprehension. There are two aspects of this experience: there is a judgment about the right thing to do and a motivation to act in certain way, and when it is disregarded it gives rise to a mental discomfort. The judgments about right and wrong are made by reason but we still refer to conscience as a final arbiter to discern right from wrong. So, there must be a difference between the judgments of practical reason and conscience. What is distinctive about conscience is, its authority. It creates a sense of obligation that it is stronger than the conclusions of practical reason. So the voice of conscience is actually the voice of an authority that has a motivational force to direct us reflect on our actions and change them if necessary.

What is the source of this authority? The philosophers, which have been examined in this study give different answers to this question. I argue that an examination of their answers in terms of their content and function will contribute to our understanding of conscience.

The term conscience emerged around 5th century BC among the ancient Greeks. The original Greek term *suneidēsis* (συνείδησις) indicates the kind of knowledge that is shared with others or with oneself, but soon it acquired the specific association of being a potential witness for or against the person with whom the knowledge is shared. Eventually, the content of conscience becomes as one’s knowledge of one’s own bad deeds and following disturbing feelings.
I argue that Plato and Paul’s contribution to the conceptualization of conscience provide the framework of the subsequent developments of the term. Plato used the term in his dialogues without elaborating on it. However, he uses the metaphor of a close relative of his (a part of his self), who lives in the same house with him (in the same body) and he hears every insult from that man for his misdeeds (the voice of conscience) (*Hippias Major*, 304c-e). The authority of the “close relative” comes from his wish to maintain the harmony of the soul. And the function of his voice is to compel us to examine ourselves in terms of our deeds, thoughts, beliefs and values. Whenever we engage in a silent dialogue with ourselves there is a possibility of encountering a contradiction within ourselves which can distort the harmony of our soul. The fear of having a disharmonious soul and a blaming and insulting relative at home is what conscience is. Plato’s understanding of “conscience” requires that one has to care for her soul in order to hear the voice of conscience. This is also necessary to build the content of conscience. The truth about right and wrong is gained through constant care of the soul. One of the ways to take care of our soul is to employ the method of elenchus on ourselves, so that we can be aware of our false beliefs and inconsistencies and eliminate them. Only someone, who feels accountable for her deeds, can engage this kind of investigation. Hence, Plato’s main contribution to the conceptualization of conscience is his creation of a role model, i.e., an ideal person, who thinks that a part of herself is watching and accusing her according to the criteria she has committed because she feels accountable herself for her deeds.

Paul’s main influence on the development of conscience is his introduction of the idea that conscience works on the basis of the law written in the heart of men. The main
function of conscience to bear witness, accusing or excusing one on the basis of this law on judgment day. Remarkably, Paul claims that although the law written in the heart of men is infallible, one’s understanding of it may be fallible. So the authority of the conscience does not originate from the law that is given by God, but from its being an agency that gives witness to God.

Medieval thinkers are influenced by Paul’s association of conscience with the God-given inner law of right and wrong. The main problem for them is that given the infallible nature of the inner law, how is it possible to sin? They attempt to solve this problem by analyzing conscience in terms of its two aspects: synderesis, the content, and conscientia, the function. Synderesis, as the infallible source of the knowledge of right and wrong, provides the natural precepts and also it murmurs back to us in answer to sin. Whereas, conscientia deliberates and draws particular conclusions for action. The relation between synderesis and conscientia is a complicated one and the importance of medieval thinkers lies not in the explanations they give but in the questions they ask regarding their nature and relationship: Is synderesis a disposition or a potentiality, if it is a potentiality then does it belong to reason or desire? Similarly, is conscientia a disposition or a potentiality, if it is a potentiality then does it belong to reason or desire? How does synderesis murmur back in answer to sin? These questions attempt to establish that there is a distinct entity in the soul, whose function is different than other functions of reason and desire. For example, when Philip the Chancellor and Thomas Aquinas regard synderesis as an innate disposition of reason, they presume that general principles of natural law are known to human beings without further reasoning. Similarly, when Thomas claims that conscientia is the act of applying these universal principles to a
particular situation, he differentiates *conscientia* from syllogism. Although these answers are dubious, as discussed in the second chapter, they reveal the need for such a clarification, to establish conscience as a distinct entity.

Another merit of the examinations of medieval discussions on conscience is to point out a possible source of the authority of conscience. The basic assumption of this approach is that we have to listen to the voice of conscience because it originates from the law of God. However, these laws are so general that they need to be interpreted by the fallible *conscientia*, most of the time. So, when we hear the voice of conscience, does it really come from God? Probably, but not directly. None of the medieval thinkers think that conscience makes us infallible; human beings are fallible, even though they insist that *synderesis* is infallible. What makes *synderesis* infallible is not the abstract content it holds but its being directed toward good. Therefore, the authority of *synderesis* originates from its call for the good for the sake of being accountable before God.

Kant dwells on both merits and pitfalls of the medieval understanding of conscience and rehabilitates it. His main contribution to the tradition is that he provides coherent answers to the questions about the distinction between reason and conscience. According to Kant, conscience is a “sensuous pre-concept” (*Ästhetische Vorbegriff*), it is one of the preconditions of being an ethical agent because it is “consciousness of what duty is” (Kant. Ak. 27:614). He conveys most of the functions of *synderesis* and *conscientia* to practical reason and judgment. The source of moral knowledge is the practical reason. Understanding provides rules, which determine objectively the rightness or wrongness of an act and then judgment brings cases under laws. Hence, in the Kantian view, conscience does not have a content and it does not carry out the reasoning. With the
consciousness of duty, it calls reason to scrutinize the moral judgments and witness the process of the trial to check whether it is conducted properly or not. And then conscience puts the verdict into force: acquitting or condemning. Hence, Kantian conscience still follows the traditional account, murmuring back in answer to sin/wrong-doing and directing one to do the right thing.

Kant not only re-secularizes conscience, which emerged as a secular concept originally, by introducing the universal law, but also he revisits the idea of an ethical person of Plato (Gorgias 478d-480d; Hippias Major 304c-e) and Aristotle (Nicomachean Ethics 1166a), who scrutinizes himself for his own sake, to be accountable for himself:

This command is “know (scrutinize, fathom) yourself,” not in terms of your natural perfection (your fitness or unfitness for all sorts of discretionary or even commanded ends) but rather in terms of your moral perfection in relation to your duty. That is, know your heart – whether it is good or evil, whether the source of your actions is pure or impure, and what can be imputed to you as belonging originally to the substance of a human being or as derived (acquired or developed) and belonging to your moral condition.

Moral cognition of oneself, which seeks to penetrate into the depths (the abyss) of one’s heart which are quite difficult to fathom, is the beginning of all human wisdom. For in the case of a human being, the ultimate wisdom, which consists in the harmony of a being’s will with its final end, requires him first to remove the obstacle within (an evil will actually present in him) and then to develop the original predisposition to a good will within him, which can never be lost. (Only the descent into the hell of self-cognition can pave the way to godliness.) (Kant. Ak. 6:441, emphasis added)

Hence, it is possible to conclude that in order to make oneself accountable, one has to scrutinize one’s own deeds, beliefs and eliminate any inconsistencies. The function of conscience is to call for a self-examination and at the end of the examination, it

---

When the term conscience emerges it is not explained with reference to the gods. For this reason I consider it as a secular concept originally, though this use is anachronistic.
pronounces the sentence: acquittal or conviction. And the conviction is followed by guilty feelings and remorse.

Nietzsche is depicted as one of the main opponents of this view. I find this evaluation correct for the most part, especially when we take his view on bad conscience into consideration. Nietzsche attacks the morality of *mores* and the internalization of all kind of values, which are imposed on human beings during civilization, including conscience as a consciousness of guilt before god and society. He develops a genealogy of conscience, showing that bad conscience originates from turning the cruel instincts inwards as a result of civilization. He claims that by means of inflicting pain for one’s debt, human beings make this inner suffering meaningful. Human beings are guilty before god and society because they owe their existence to the god and the society and they are not able to repay this debt, moreover it is not possible to redeem the original sin, i.e., human beings are a total failure because of their faulty nature. Bad conscience is the consciousness of this guilty feeling, which is considered by Nietzsche as a sickness. Healing is possible when human beings get rid of the idea of god, the ultimate source of guilty feelings. Only then human beings can become sovereign individuals with a sovereign conscience that is “the proud awareness of the extraordinary privilege of *responsibility*, the consciousness of this rare freedom, this power over oneself and over fate” (GM II, 2).

The merit of Nietzsche’s genealogy of conscience neither arises from the accuracy of the historical account, which is dubious for the most part, nor from affirming Kant’s views on conscience – consciousness of responsibility/duty76 – but from providing an

---

76 Obviously they differ completely on what is duty.
insight about the content of conscience. So that, the content of conscience may not be
some necessary truths about the right thing to do but may come from contingent beliefs,
which are imposed on us by the society in which we are brought up. If this is the case
then there would be some consequences of this claim. The verdicts of conscience could
be at most inter-subjective or in the case of Nietzsche’s sovereign individual, utterly
subjective. Moreover, we observe that people who rely on conscience often approve of
radically different practices, including some that may seem outrageous. And if we recall
how often cruel and destructive conduct has been excused in the name of conscience,
then it seems questionable whether conscience is each individual’s unerring access to
ethical truth. Given this doubt about conscience it is legitimate to ask further questions: to
what extent and why should one (i) consider the promptings of one’s own conscience as
one’s authoritative guide and (ii) respectfully tolerate the conduct of others when they are
apparently guided by conscience?

Given the findings of this study it may be concluded that the authority of
conscience arises neither from the content, whatever they be, nor from the ways it derives
conclusions but from the belief that I could not live with myself if I do something wrong,
so I have to scrutinize my actions and judgments: if I find something wrong about my
past action I have to admit the guilt or if I find something wrong about my future action, I
should not do it. This belief originates from one’s awareness of one’s accountability for
one’s actions as one of the most common usages of the original term: σύνοιδα ἔμαυτῷ: I
know with myself that; I bear witness to my conduct that. Actually, having an awareness
for one’s accountability for one’s actions is a pre-requisite of being an ethical person.
Everybody can follow the mores of her own community and can lead a “moral life”, but
being an ethical person requires being on call for scrutiny of our actions to check whether they are in line with our ethical commitments.

In accordance with the claims that are examined in this study, we can say that conscience is infallible as an act of submitting our actions and decision for further examination, but it is fallible regarding the results of this examination, because it is conducted by reason and reason may be erroneous.

On the basis of this understanding of conscience, it is possible to answer the questions above: (i) The promptings of one’s own conscience can be one’s authoritative guide because although they are fallible, they arise from one’s sincere wish to be in line with her ethical commitments. (ii) In order to respectfully tolerate the conduct of others when they are apparently guided by conscience we need the evidence that the person who claimed to have a clear conscience is able to show that his actions or judgments are in line with his ethical commitments. For example, when Adolph Eichmann claims that he has a clear conscience, he has to show that sending Jewish people to the concentration camps is compatible with the principle “it is wrong to kill people”. Obviously he would fail to fulfill this task. Then how is it possible for him to claim that he has a clear conscience? This is due to his belief that he is not responsible for these killings and hence he cannot be held accountable for these wrong-doings. Without admitting the accountability or becoming aware of one’s accountability it is not possible to exercise conscience. These examples confirm the attempts to define conscience primarily as the consciousness of one’s own duty/accountability.

---

77 In his trial he testifies that he believes in God. On the basis of this testimony it will be legitimate to assume that he commits to this principle. See Arendt 1963, Eichmann in Jerusalem and The Trial of Adolph Eichmann: Complete Transcripts retrieved at: http://www.nizkor.org/hweb/people/e/eichmann-adolf/transcripts/Sessions/index-03.html
Following this discussion, I argue that conscience can have a subjective authority for the ethical agent herself but we cannot attribute an objective authority to conscience. For this reason, when we accuse someone for not having a conscience, actually, we blame this person for not acknowledging his responsibility for a wrong-doing or/and for an erroneous conclusions of his ethical reasoning. Similarly, when we say that we appeal to conscience, we mean that we submit our ethical judgments to further investigation but not to mean that we appeal to an objective authority about the right and wrong.

If we go back to the initial question of this study regarding the ambiguity of conscience, I believe that the findings of this study enable us to have a clearer understanding of conscience in terms of its content and function. Given this understanding, what can be the significance of conscience for us in ethical matters?

Following Kant, I suggest that conscience needs to be repeatedly awakened, with frequent evocation of consciousness of ones deeds. One must also sharpen attentiveness to the voice of conscience and use every means to obtain a hearing for it. So that, we can be more aware of our accountability and responsibility in our deeds and we can test our sincerity regarding our ethical commitments in particular cases.
Bibliography


