ADDIS ABABA UNIVERSITY
SCHOOL OF GRADUATE STUDIES
FACULTY OF SOCIAL SCIENCIES
DEPARTMENT OF PHILOSOPHY

A BERIEF ANALYSIS OF IMMANUEL KANT'S ETHICAL SYSTEM: THE CATEGORICAL IMPERATIVE

BY

ABRAHAM TSEHAY

July 2014
Addis Ababa

ADDIS ABABA UNIVERSITY
SCHOOL OF GRADUATE STUDIES
FACULTY OF SOCIAL SCIENCIES
DEPARTMENT OF PHILOSOPHY

A BERIEF ANALYSIS OF IMMANUEL KANT'S ETHICAL SYSTEM: THE CATEGORICAL IMPERATIVE

BY

ABRAHAM TSEHAY

ADVISOR

TENNA DEWO (PhD)

A THESIS SUBMITTED TO THE DEPARTMENT OF PHILOSOPHY
SCHOOL OF GRADUATE STUDIES
ADDIS ABABA UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENT FOR THE DEGREE OF MASTER OF ARTS IN PHILOSOPHY

July 2014
Addis Ababa

ADDIS ABABA UNIVERSITY
SCHOOL OF GRADUATE STUDIES
FACULTY OF SOCIAL SCIENCIES
DEPARTMENT OF PHILOSOPHY

A BERIEF ANALYSIS OF IMMANUEL KANT'S ETHICAL SYSTEM: THE CATEGORICAL IMPERATIVE

BY

ABRAHAM TSEHAY

APPROVED BY BOARD OF EXAMINERS:

Advisor	**Signature**	**Date**
Examiner	**Signature**	**Date**
Examiner	**Signature**	**Date**

July 2014
Addis Ababa

Acknowledgement

I wish to express my heartily gratitude to my advisor Tenna Dewo (PhD) for his insightful comments that helped me to make this thesis much better than it would have otherwise been. Many thanks are due to Addis Ababa University particularly the Department of Philosophy for giving me a chance to study in this area. I also want to thank my family for the encouragement they have given me to finish this thesis. Last but not least, I would like to thank all of my friends whose comments helped me a lot to improve this thesis.

Abraham Tsehay

Table of contents Page numbers

Notes on Sources and Key to Abbreviations

Works by Immanuel Kant

There is unfortunately no generally accepted method for referring to passages in Kant's work. In this thesis, apart from the *Critique of Pure Reason*, where references are to the standard A and B pagination, all citations refer to the following English translations which include the volume and page number of the Prussian Academy edition of Kant's work.

G Kant, Immanuel. 1998. *Groundwork of the Metaphysics of Morals*. Translated and edited by Mary Gregor. Oxford: Oxford University Press.

Kant, Immanuel. 1993. *Grounding for the Metaphysics of Morals: With on a Supposed Right to Lie because of Philanthropic Concerns*, trans. James W. Ellington. Indianapolis: Hackett publishing company.

Fi Kant, Immanuel. 200. *First Introduction to the Critique of the Power of Judgment*. Translated by Paul Guyer. Cambridge: Cambridge University Press.

CPR Kant, Immanuel. 1998. *Critique of Pure Reason*. Translated and Edited by Paul Guyer and Allen Wood. Cambridge: Cambridge University Press. Cited by A/B pagination.

cpr Kant, Immanuel. 2002. *The Critique of Practical Reason*. Translated by Werner S. Pluhar. Indianapolis/Cambridge: Hackett publishing company, Inc.

LE Kant, Immanuel.1997. *Lecture on Ethics*. Translated by Peter Heath. Cambridge: Cambridge University Press.

Mm Kant, Immanuel. 1991. *The Metaphysics of Morals*. Translated by Mary Gregor. New York: Cambridge University press.

Formulas

FUL The Formula of Universal Law: "Act only according to that maxim through which you can at the same time will that it become a universal law" (G 4:421)

FLN The Formula of the Law of Nature: "So act as if the maxim of your action were to become by your will a universal law of nature" (G 4:421).

FH The Formula of Humanity as an End in Itself: "So act that you use humanity, in your person as well as in the person of any other, always at the same time as an end, never merely as a means" (G 4:429).

FA The formula of autonomy "*act in accordance with a maxim that can at the same time make itself a universal law"* (G 4:437).

FRE the formula of the realm of ends "*Act in accordance with the maxims of a member giving universal laws for a merely possible realm of ends (*G 4:439)."

Introduction

Immanuel Kant (1724-1804), a German philosopher, is considered as the father of modern ethics and one of the great philosophers in the history of philosophy. He wanted to lay unshakable foundation for moral philosophy. His focus on moral philosophy came after his work entitled the *Critique of Pure Reason,* which served as the foundation for his practical reason. His contribution to ethics is totally new which was not attempted by earlier ethicists. He wanted to show, by using reason, that ethics and morality are based on a single supreme universal principle, which is binding to all rational beings. In precise words, Kant wanted to establish the first principle of morality which neglect all consideration of self-interest and even particular human problems (Paton 1947, 15).

We can find Kant's most influential position in his works, *The Groundwork of the Metaphysics of Morals (*we call it *Groundwork* hereafter*)* and developed his views in his later works such as *The Critique of Practical Reason, The Metaphysics of Morals,* and the like. In this Thesis, I will mainly concentrate on the foundational doctrine of the *Groundwork.* I will also consider some of his other books like the *Critique of Pure Reason,* the *Critique of Practical Reason*, the *Metaphysics of Morals* and others in case situations impel me.

In the *Preface* of the *Groundwork*, Kant claims that his intention is to seek out and establish the supreme principle of morality, and that supreme principle is the categorical imperative. This is, Kant argues, one and only one supreme principle for all fields of morals (G 4:392).

However, Kant does not insist that he discovered categorical imperative. Because he thinks that this supreme principle is ordinarily presumed in all moral judgments. That is, it is a working criterion used by any rational agent to make their choice and judgment although they do not explicitly formulate it (Ellington 1993, vi). In other words, he aimed at exposing the moral principle

which is implicitly used by the pre philosophical understanding of morality. He insists that although in pre philosophical understanding of morality any ordinary good man used this moral principle to judge moral actions, he/she was unable to make these principles clear for themselves. They were also unable to clearly differentiate it from other principles focusing on happiness of the individual or benefits which comes from moral life (Alison 2011, 71). Thus, in section one of the *Groundwork,* Kant starts from the analysis of our ordinary moral views in order to determine the supreme principle behind it. As a result, he argues that the only thing good without qualification is a good will. The good will is good through its willing; that is, only a good will and actions that express this will have this special kind of value, namely unrestricted or unconditional value (G 4:393).

To ascribe unconditional value to an action, for Kant, we have to know the motivation on the basis of which a person acts. Thus, what makes us attribute unconditional value to a morally good action is the motivation behind it, or the principle upon which it is willed. This suggests that if we know how actions with unconditional value are willed, we can know what makes them morally good. If we know what makes actions morally good, we can determine which one is morally good and what the moral law tells us to do. In order to determine what the moral law tells us to do, Kant attempts to find out the principle on which a person of a good will acts (Korsgaard 1994, xii).

According to Kant, a person of a good will performs out of the motive of duty. Consequently, if we analyze the concept of duty or actions done out of the motive of duty, we can show what the principle of a good will is. For Kant, an action motivated out of duty is not performed for the sake of some purpose, such as helping others who are in need. The defining feature of an action performed from duty is that the agent performs the action because he/she considers that it is the right thing to do. In other words, for the agent doing the action or fulfilling its purpose is

something expected of him. This defining feature of actions done from duty lies in the *maxim* (subjective principle*)* upon which it is done or upon which the agent performs the action (G 4:399). Kant holds that an agent who acts from duty or from good will considers his/her maxim as having the form of a law. The principle of a good will is, thus, to perform only actions whose maxim can be conceived as having the form of a law (Korsgaard 1994, xv).

In section two of the *Groundwork,* Kant arrives at the same conclusion but through different way. He does not start from the analysis of common moral cognition but from the general analysis of the role of reason for action. He chooses a philosophical starting point that leads him to a more complete and precise formulation of the moral principle. He analyses the concept of unconditional necessary action or the concept of obligation to uncover a principle similar to section one. Unconditionally necessary actions are moral actions. If an action is unconditionally necessary, the agent considers that doing the action is something required of him even if there are encounters from inclination or self-interest. Thus, the maxim upon which it is done should be conceived as a law. But, since the action is unconditional, the will cannot be bound to a particular law. As a result, it is the idea of law that is universal that binds the will (G 4:412-21). This takes us to Kant's first formulation of categorical imperative, formula of universal law (FUL): "Act only according to that maxim through which you can at the same time will that it becomes a universal law (G 4:421)".

For Kant, to determine whether one wills his/her maxim to be a universal law, he/ she can consistently will it as a law of nature. Nevertheless, the kind of contradiction that he thinks in his mind is controversial because it is not clear and it seems to differ from example to example that he provides (Korsgaard 1996, 77-78).

Analysis of the concept of moral obligation (imperative) shows that the FUL is its principle. But, it is impossible to justify it by analyzing the concept of moral obligation; it is impossible

to show how it is binding on the will through mere analysis of the concept of moral obligation. Kant postpones this job to section three of the *Groundwork* (Paton 1948, 30). In preparation to resolve this problem in section three of the *Groundwork,* Kant has to show the kind of motivation that an agent who acts from categorical imperative has. He insists that since every action has an end, morality become possible if there exists an objectively necessary end. Kant here introduces humanity as the only end which is an end in itself. This enables him to formulate another version of the categorical imperative, namely the formula of humanity (FH) (G 4:429).

Kant comes up with another formula of the categorical imperative by combining the previous two formulas (FUL and FH). He takes the idea of legislation from FUL and the idea of self-determination from the FH to establish autonomous legislation in the kingdom (realm) of ends. For Kant, one might be motivated to obey the law either heteronomously or autonomously. Since a person who acts heteronomously is motivated to respect the law by some interest, the imperative from which his/her maxim is drawn is hypothetical. However, the moral imperative is categorical. A person who is motivated by it, acts autonomously. This implies that the laws of morality must be laws that a person imposes on his/herself. Kant suggests that any autonomous action must be governed by moral laws. The moral law just tells a person not to act on a principle that he/she does not will to be a law. It restricts a person to act in accordance with his/her autonomy. As a result, Kant suggests that the categorical imperative is the law of autonomy (Timmermann 2007, 102).

In section one and two of the *Groundwork,* Kant only shows, through analysis, the principle of a good will and the kind of motivation involved in acing from it. In other words, in section one and two of the *Groundwork,* he tells us what morality is if it really exists. He does not tell us that morality really exists or that the categorical imperative has a binding force on us. Kant shows us this in section three of the *Groundwork.* To do so, he uses a synthetic rather than analytic argument.

Kant insists that, as rational beings, we must consider ourselves as possessing a free will. We cannot consider our actions as causally determined by outside forces; our will is independent of external influences. This is a negative conception of freedom. Thus, it is uninformative. However, this negative concept of freedom is important to know the positive one. Kant leads to the positive concept of freedom in such a way that a lawless free will is self-contradictory. Thus, a free will would act under laws, but these laws could not be one imposed upon it other than itself. Because if so, they would be just laws of natural necessity. If the laws under which free will act upon are not other imposed, they must be self-imposed. But, this is what does mean by autonomy. Since, for Kant, autonomy is the principle of morality, a free will is under moral laws. Therefore, morality follows from the concept of free will (Paton 1948, 41).

Since morality is valid for all rational beings, it is impossible to base our notion of morality on the concept of free will if it is not possible to prove that all rational beings have free will. It is impossible to prove this issue from experience. However, it is possible to assume that a being is really free if it thinks of itself as free when it acts. This is because a being with reason and a will must think of itself as free; if reason were influenced by external forces outside of itself, it would not be reason at all (G 4: 448).

However, Kant cannot be satisfied by this argument because it is circular. He attempts to avoid this problem by establishing an independent ground to consider the will as free. The distinction between the noumenal and phenomenal world enables him to establish this independent ground. He insists that the will is part of both the noumenal (intelligible world) and the phenomenal world (the world of appearance). He says it is only in the phenomenal world that human being is causally determined. But, as parts of the noumenal world, the will is free (G 451-53).

As I have tried to make clear before, Kant's aim in the *Groundwork* is to search for and establish the supreme principle of morality (i.e., categorical imperative). He attempted to do this in section three of the *Groundwork* in which he affirmed that the noumenal world makes freedom possible since the idea of freedom makes any moral agent to be responsible for his/her act; moral obligation is real. But, to me, the way he attempted to justify a categorical imperative is problematic. Thus, in this paper, I argue that Kant does not put the categorical imperative or morality on a solid ground.

This paper is composed of three chapters each of which deals with a specific part of the thesis. In chapter one, I will briefly discuss and analyze how Kant, in the *Groundwork*, arrives at the first formulation of the supreme principle of morality by analyzing ordinary moral beliefs. In chapter two, I will analyze how Kant reaches the *five* or *three formulation*[1] of the categorical imperative by taking more philosophical stand point (by analyzing the concept of rational will or rational agent). In chapter three, I will analyze how Kant attempted to establish the notion of categorical imperative in the *Groundwork* and I will also show how his attempt is unsuccessful. Finally, in conclusion part, I will briefly show how the reason he used to establish the validity of the categorical imperative, in the *Groundwork*, is insufficient.

[1] The number of formulas for the categorical imperative that Kant gives are controversial. In fact, Kant explicitly states only three formulas (G 4:436). In spite of this, Guyer (1998, 216) maintains that there are four, Allison (2011, 150) maintains that there are three, Paton (1947, 129-32) claims that there are five. The appearance of five is suggested by the fact that both the first and third have variants or comes in two forms.

Chapter One: Approach to a Categorical Imperative

1. The Good Will and the Supreme Principle of Morality

1.1. Overview

The "*Groundwork*" is divided into three sections. In each section, Kant accomplishes a specific project which in turn contributes to the fulfilment of the whole project (Korsgaard 1994, xi). In section one of the *Groundwork*, his project is "to proceed analytically from common moral cognition[2] to the determination of its supreme principle (G 4:392)". He starts from our ordinary ways of thinking about morality and moves into discovering the principle behind it (Korsgaard 1994). Christine M. Korsgaard reminds us that although Kant, in this section, is analyzing[3] ordinary moral view, he is not justifying that human beings have moral obligation. He is rather merely identifying what is essential to prove that moral obligation is real (Korsgaard 1994, xi).

1.2. The Good Will

"The common cognition from which Kant starts his argument is that morally good actions have a special kind of value. A person who does the right thing for the right reason evinces what Kant calls a good will (Korsgaard 1994, xi)." Thus, Kant's analysis of ordinary views of morality

[2] According to Allen W. Wood, the term "common moral cognition refers to an everyday unreflective awareness of the rational standards Kant thinks anyone must use in moral deliberation and judgment (Wood 1999, 20-21)".

[3] By "proceed analytically" Kant means that we move from what we suppose we know and discovers what that knowledge presupposes.

begins with the thought that “It is not possible to think of anything in the world, or indeed out of it, that can be held to be good without limitation except a good will (G 4: 393)”.

Here, the objective is to identify or state the principle that ordinary good man is supposed to use in judging his action although he does not explicitly formulate it. Kant’s intention is to derive it by the analysis of the concept of a good will. Thus, to explicitly state this principle, we have to know more about “a good will” and “good without qualification” (Paton 1947, 34).

But, before I proceed to the discussion of the concept of a “good will” and “good without limitation”, I find it important to consider the phrases “not possible to think” and “out of the world”. According to Hennery E. Allison, the first phrase shows that Kant does not just claim that we cannot find anything which is “good without limitation except a good will”. We cannot even conceive of its possibility. Whereas the latter phrase shows Kant’s two intentions. The first and the apparent intention is to show that the concept of a good will is “normative” for both holly will and human will. The second intention (although not as apparent as the first one) is to show other possible world to which the unique goodness of this concept is supposed to be applicable (Allison 2011, 71-72).

By “good without limitation”, Kant wants to mean that it is only a good will which is good in any context. It cannot be good in one situation and bad in another situation. It cannot be good as a means to one end and bad as means to another. The goodness which is attributed to a good will is not also subject to individuals’ desire. From this, it is quite possible to understand that its goodness does not change along with the change of contexts, desires and ends. As a result, Kant concludes that good will is an absolute and unconditioned good. It is also the only thing which is good in itself and independently of other things (Allison 2011, 72-73).

Kant here does not mean that a good will is the only good. Thus, he analyses other goods to support the above idea. He divides all goods into two: "gifts of nature" and "gifts of fortune". "Gifts of nature" is further divided into "talents of the mind" and "qualities of temperament". Talents of the mind comprises of qualities such as understanding, wit and judgment, whereas temperament includes qualities such as courage, resoluteness and persistence. Kant gives us power, wealth, honor, health and happiness as examples of "gifts of fortune". Although Kant does not deny that all of these goods are good in many respects, he does not confirm that they are good without limitation. He claims that, on the contrary, they can be extremely bad or evil when they are used with a bad will. This implies that they are good within certain conditions (only when they are used with a good will); or they are good only when some conditions are fulfilled. Thus, they are conditioned good and are not absolutely good or good in themselves (Paton 1948, 17).

Moreover, Kant believes that the goodness of the good will is not derived from the good result it produces; it does not add anything to its inner unconditioned or absolute value. But, it just attaches a distinct, conditioned and qualified good that does not have any influence on its inner nature. Sometimes, a good will might be combined with defects of the mind or bad qualities of temperament. In this case, it may fail to achieve the good it wills or fails to achieve good results. But, this bad result takes nothing away from it; that is, a good will continues to have its inner unconditioned or absolute value although, due to some misfortune, it fails to achieve the result it aims to achieve (Wood 1999, 21-22).

According to H. J. Paton, it does not, for Kant, mean that a good will does not aim at producing results. On the contrary, Kant believes that a good will and any kinds of will must aim at bringing results (Paton 1948, 17).

Ordinary moral consciousness agrees with the view that it is only a good will that is unconditionally good. Indeed this is the condition or foundation of all ordinary moral judgments (G 4:394). However, Kant thinks that "there is something so strange in the absolute worth of a mere will without any consideration of utility in its estimation (G 4:394)" which may arise "the suspicion that such an idea may perhaps have as its hidden basis merely some high-flown fancy (G 4:394)." Kant now attempts to defend his idea that moral goodness is the primary purpose of reason. To do so, he uses an argument that nature has the purpose of 'assigning reason' to the human will as its sovereign. Kant's argument depends on the principle of traditional teleology that nature does no work without a purpose or function. If happiness, not morally good willing, were the greatest and unconditioned good, practical reason would be superfluous since in principle instinct would be a more appropriate tool for making us happy. However, nature, Kant argues, has given us practical reason (a will); and it must have happened for a purpose or function. If its function is not to make us happy, it must be to produce a will which is good in itself (G 4:396).

1.3. A Good Will and Duty

Kant does not discuss the concept of duty independent of other concepts. He discusses it to further clarify the concept of a good will. His main claim is that the concept of duty "contains that of the concept of a good will though under subjective limitation and hindrances, which far from concealing it and make it unrecognizable, bring out by contrast and make it shine forth all the more brightly"(G 4:397). This passage is incorrectly interpreted by many readers of the *Groundwork*, but Kant's intention in this passage is that a will which acts for the sake of duty is a good will. However, as H. J. Paton suggests, it should not follow from this that a good will is necessarily one that acts for the sake of duty. To understand this, we have to make clear the concept of duty. The very idea of duty involves the overcoming of desires and inclinations. On the other

hand, a completely good will or "holy" will can manifest itself without overcoming natural inclinations and desires. So, a good will as such would not act for the sake of duty since it includes a holy or a perfectly good will that manifests itself without being constrained by duty. However, man's will is not holy or completely good since in finite creatures [such as man], there are certain "subjective limitations" or obstacles. Thus, a will that acts for the sake of duty is a good will under human conditions (Paton 1947, 46). This also implies that the concept of a good will as such is broader than the concept of a will that acts for the sake of duty (a good will under human conditions). Thus, a will which acts for the sake of duty is a special form of a good will that overcomes subjective limitations and hindrances. It manifests itself in duty under adverse situations. According to Allen W. Wood, Kant takes us to this special form of a good will, because the good will's unlimited worth is shown more brightly to common rational cognition in this conditions. Wood also reminds us that Kant directs our attention to this special form of a good will because it is under this condition the fact that a good will is different from other aforementioned goods such as qualities of temperament, and then its higher value becomes clear (Wood 1999, 27).

1.4. Elucidation of the Concept of Duty by the three Propositions

1.4.1. Moral Worth

Kant elucidates the concept of duty by three propositions of which he explicitly states only the second and the third propositions.[4] The most natural or appropriate candidate of the first proposition, although he does not explicitly state it, seems like: *an action has moral worth or value only when it is done from duty* (Paton 1948, 18-20) and then Kant compares it with actions that

[4] Kant does not explicitly state the first proposition about duty. He explicitly states only the second and the third proposition. But, his suggestion at (G 4:418) that the third proposition is the consequence of the last two propositions impels commentators to search for the first proposition.

conform to duty. An action is said to conform to duty if it complies with what duty necessitates or requires regardless of the motivation for doing it. Kant ascribes true "moral worth" or "moral content" only to actions done from duty (G 4:397-98). The term "moral worth" or "moral content", as Allen W. Wood reminds us, "do not refer to just any sort of value morality might attach to actions, but designate only that special degree of worth that most conspicuously elicits *esteem* from common rational cognition"(Wood 1999, 28). However, if an action lacks moral worth in this sense, it does not follow that it is worthless from moral point of view. In fact, all actions that conform to duty have value from moral point of view, but they do not just have a distinctive "moral worth"; that is, they do not just have a special degree of value that goes beyond such mere moral approval and elicits *esteem* from common rational cognition or elicits a recognition of unconditional value that deserves only to a good will (Wood 1999, 28). Wood argues that if Kant really held that actions done in conformity with (but not from) duty reflected an immoral or morally worthless will, this would be self-contradictory. Because this could imply that actions which conform to duty are wrong or immoral. If so, they cannot conform to duty after all (Wood 2008, 33).

In the case of actions that conform to duty, Kant identifies two main kinds of actions. The first action is one done from self-interest. These are actions done as a means to satisfy a higher order inclination directed to another object. An action done merely out of self-interest might conform to duty. In this case, it might be right, but it does not have special degree and distinctive moral worth. As an example of this, Kant gives us a shopkeeper who does not overcharge his inexperienced customers not from an immediate inclination such as sympathy (love) to treat them honestly, but from calculating a long run profit. He is, in short, the one who considers honesty as the best policy to maximize his profit in the long run. In as much as this salesman consistently treats his customers honestly, his act conforms to duty. In this case, the shopkeeper's act is right,

but it does not have a genuine moral worth (G 4:397). The reason is that the shopkeeper does not present us with unambiguous case of action from duty. If we assume that, as Kant expects us, a long run self-interest motivates him to serve all his customers honestly, we have little reason to accept his action is done from duty while observing the shopkeeper serving his gullible or inexperienced customers honestly. Moreover, the nature of morality becomes apparent when it conflicts with inclination, and in the case of the prudent salesman morality and inclination simultaneously occur (his commercial interest invariably recommends what duty commands). Thus, his action is uninformative and his mettle cannot be put to test. This also implies that Kant is skeptical about the actual moral attitudes of human being (Timmermann 2007, 31).

The other kind of actions that conform to duty are those that are done from immediate inclination (actions done for their own sake rather than for some extrinsic end). For instance, beneficence from immediate sympathy are actions of this type. We may incline to ascribe moral goodness to this action. To test this, according to H. J. Paton, we must isolate our motives: first we have to isolate actions done only from the motive of inclination, and second we have to isolate actions done only from the motive of duty. If we do so, we can find that actions done from immediate inclination may be praise worthy and right but they do not have a distinctively moral worth. However, the same action done only from the motive of duty has a distinctively moral worth (Paton 1948, 19).

Timmermann insists that "Kant does not wish to say that actions possess moral worth only when they are done without any concurrent inclinations, or even contrary to opposing inclinations; even less that these adverse conditions are desirable or that we should try to bring them about". But, Kant says that only action that is not supported by inclination, or even resists strong inclination possesses moral worth, because the moral content of the action becomes certain or apparent

only in difficult cases (Timmermann 2007, 33). In this sense, the existence of inclination does not detract from the moral worth of the action, so long as the action is solely determined from the motive of duty. Otherwise Kant's argument would be absurd (Paton 1948, 19). The only thing that has to be certain is that the action is not determined by unnecessary motives of inclination (Timmermann 2007, 33).

Moreover, Kant does not reject happiness altogether. But, rather, he holds that we have at least an indirect duty to seek our own happiness. This is because "discontent with ones condition under many pressing cares and amid unsatisfied wants might easily become a great temptation to transgress one's duties" (G 4:399). Thus, an indirect duty, as Timmermann says, is not a duty in the proper sense. That means, it is not morally binding by itself. It is not a kind of duty alongside with perfect and imperfect duty officially presented by Kant at section two of the *Groundwork* (Timmermann 2007, 36).

1.4.2. Moral Worth and Maxims

Kant's second proposition about duty is that: "an action from duty has its moral worth not in the purpose to be attained by it but in the maxim in accordance with which it is decided upon, and therefore does not depend upon the realization of the object of the action but merely upon the principle of volition in accordance with which the action is done without regard for any object of the faculty of desire (G 4:400)." In the first proposition, we have seen that an action has moral value only if it is done from duty. The above proposition adds that the moral value of such an action (i.e., an action done from duty) does not stem from the result that the action produces or it seeks to produce (G 4:399-400). If the moral worth of the action cannot be derived from any inclination of achieving something, then it cannot be derived from the result it is sought or produced. The result we seek to produce by our action is not necessarily identical with that we in fact produce

by it. Sometimes, a morally good action may not even have any result. H. J. Paton notes that Kant, here, dismisses all forms of utilitarianism (Paton 1947, 58).

According to H. J. Paton, Kant does not deny that actions done from duty will produce results and seeks to achieve results. They always seek to attain a result, and they succeed in producing the sought result. What he says is that the moral worth ascribed to actions does not rely on the result they seek to achieve or in fact produce (Paton 1947, 58-9).

The idea so far has been negative; the source of moral worth is not the result that the action in fact produces or aims to produce. So, it is uninformative. It is necessary to make this doctrine more positive. If an action from duty does not derive its distinct value from the result it achieves or seeks to achieve, it must more specifically be from the motive of duty. This motive must not be just a mere desire to produce some result (Paton 1947, 59). Kant expresses it in such a way that, as the above quotation about the second proposition of duty shows, an action done from duty gains its moral worth from a maxim, and the maxim is not a maxim of producing results (G 4:400). I will show how Kant makes a relation between motive and maxim below.

According to Kant, a maxim is a particular *principle* (in a loose sense) that one follows to perform a certain action. It is considered as a purely personal principle to the agent. Here, as H. J. Paton warns us, a maxim should not be confused with what is called a *copybook maxim* since it cannot be established in accordance with an already established principle while the latter has this character.[5] For Kant, it is a 'subjective' principle. It is subjective, because it is a principle that a rational agent follows to accomplish a certain action. Subjective principles are valid only for an

[5] H. J. Paton (Paton 1947, 59-61) notes that principle in general is a universal proposition which contains other propositions under it. Strictly speaking, a principle should not have a higher ground except it is grounded under the idea of reason itself; it is absolute or supreme. Kant's aim is also to discover and justify it. But, the term 'principle' can be used, in relative sense, for universal propositions which are not supreme.

agent who chooses to act upon it. In contrast, objective principles are valid for all rational agents, because it is upon them that all rational agents would choose to act only when their emotions are controlled by reason. In other words, it is a principle upon which every rational agent ought to act. A principle can be subjective only if it is a principle on which one acts. In fact, we do not always act upon objective principles. Objective principles are objective in so far as they do not depend upon the intensions of the agent. But, one may choose to act on objective principles. At this time, the principle chosen to be acted on will be both subjective and objective principle (Paton 1947, 60).

Thus, Kant's, as mentioned above, technical term for this subjective principle is a *"maxim*". It is different from an objective principle since it is not valid for all rational agents. It is also not the same as a motive, because it is more general than a motive, and that is why it is called a principle. Animals may have a motive but not a maxim. According to Kant, it is only a rational agent that can have a *maxim* (Paton 1947, 60). For instance, if one kills himself since he experiences more pain than pleasure, his maxim would be "if life offers me more pain than pleasure, I will commit suicide" (G 4:422). This is not just an impulse or a motive, but rather a general principle which determines his action and supposed to be applied in any similar condition. So, this *maxim* includes both the action and its *motiv*e. It is not, however, considered as valid for everyone else, like objective principle, and it can be good or bad (Paton 1947, 60).

All the above things give rise to a question of to what extent this subjective principle or maxim can consciously be formulated by a rational agent. However, according to H. J. Paton, Kant makes a distinction between human conduct and animal behavior, and, in acting, a human being does not blindly respond to any impulse. That means, he/she does not act without knowing what

he/she is doing; he/she also identifies the quality of his/her action; and he/she does not do this without having some concept of the principles by which he acts (Paton 1947, 60-1).

Moreover, Kant identifies two kinds of maxims: material maxim (*a posteriori* maxims) and formal maxim (*a priori* maxims). Material maxims are grounded on sensuous inclinations; they aim at fulfilling a certain desire or seek to attain a certain result. On the other hand, a formal maxim is not grounded on a sensuous inclination. It is a maxim that does not consider any particular thing. In other words, it does not aim at fulfilling a particular interest or it does not attempt to produce a particular result. It is just a maxim or principle of performing one's duty whatever that duty might be. We have seen before that a moral action, an action done from duty, does not have its moral worth either from the result that it actually produces or results that seeks to produce. If it is so, any moral action does not have its moral worth from a material maxim. It rather has its moral worth from a formal principle, a maxim or principle of performing ones duty whatever that duty might be. Thus, according to Kant, the maxim of actions performed from duty is a formal maxim rather than a material maxim (G 4:400).

1.4.3. Respect[6] for the Law

The third proposition of Kant about duty is that "duty is the necessity to act out of reverence for the law (G 4:400)". Kant holds that the maxim of a morally good action is formal maxim. If so, it must be a principle or maxim of acting reasonably. In other words, this maxim must be a maxim of acting on a law which is valid for all rational beings irrespective of their particular interests (inclinations). It is true that human beings are fallible, meaning susceptible to faults. As a result, this law must be presented in the form of duty that orders us to be obedient. Since this law

[6] Most English translators translate the German word "*Achtung"* as respect but some translators such as H. J. Paton translates it as reverence. In this thesis, I use these terms interchangeably.

is considered as an imposed one on us, it must arouse a feeling which is akin to fear. On the other hand, having realized that this law is imposed on us by our own self,[7] it must arise a feeling akin to inclination. By inclination, it is to mean a delight that results from the awareness of the imposition comes from our own free and rational will. It is this kind of feeling that Kant calls reverence or respect. This feeling does not arise from any kind of sense stimulation. It is rather from being conscious that the will is subordinated to a universal law which is free from sensuous impact. From this, one can understand that as long as a morally good action (an action done from duty) is motivated out of this kind of unique feeling, it is convincing to say that a morally good action is performed out of reverence for the law (Paton 1948, 22).

1.5. The Law

Having discussed a good will under human condition as one which acts for the sake of duty, and duty as "the necessity of actions from respect for the law" (G 4:400), the next possible question would be the nature of this law by which a good will has its unique and absolute value (G 4:402). Thus, Kant's analysis is yet to be completed since we do not still know the law that inspires reverence and motivates morally good actions, or we do not still know what the moral law exactly commands (Timmermann 2007, 44).

In describing the nature of the law by which a person of a good will performs his/her action, Kant presents us with the first version of the supreme principle of morality which is officially, for the first time, called *categorical imperative* in section two of the *Groundwork* in the following way. According to Timmermann, Kant's argument seems like this: To realize a specific end, we have to follow a certain command that enables us to use specific laws. For instance, if one intends

[7] We find here a glimpse of Kant's doctrine of autonomy officially introduced only at the end of section two of the *Groundwork.*

to make a cup of coffee, he/she has to consider the specific empirical laws that enables him/her to fulfill that end. However, we have seen before, in the analysis of the concept of duty, that morally good actions are not actions done for the sake of some end that one wants to realise or fulfill (Timmermann 2007, 44). If so, all laws used to realize a particular end cannot be candidates for being moral law. That is, these laws cannot inspire reverence and motivate morally good actions. Thus, if these specific laws are discarded, the only possible candidate that remains for the position of the moral law must be law abidingness or "the universal conformity to law as such". Kant argues that if there is something by the name moral obligation, then we have to recognize that our wills are directed by this principle: "*I ought never to act except in such a way that I could also will that my maxim should become a universal law*" (G 4:402). "The necessity of an action from reverence for this law is duty, the condition of a human will that is essentially good beyond everything else. Thus, the concept of the good will is connected with that of the law through the concept of duty (Timmermann 2007, 46)."

Kant claims that his analysis fully agrees with ordinary human reason. To state more specifically, although the ordinary good man does not establish the above moral principle in abstraction, he/she really uses it in evaluating particular moral matters. He continues to assert that even in practical affairs ordinary human reason is an important guide than philosophy (G 4:404). After showing how the principle of morality is in the possession of common human reason, Kant poses the following rhetorical question: "Might it not then be advisable to leave moral questions to the ordinary man and to regard moral philosophy as the occupation (or the game) of the philosophical specialist? (Paton 1948, 23)"

1.6. The Need for Philosophy

As noted above, Kant's optimistic assessment of the power of ordinary human reason in judging the moral worth of individual agents and their action casts doubt about the usefulness of moral philosophy. This doubt comes out of the tension between what Kant has said above and in the preface as well as at the beginning of section two of the *Groundwork* about the absolute importance of moral theory and metaphysics of morals (pure moral philosophy) (Alison 2011, 142). Kant resolves the tension in such a way that the ordinary man needs philosophy since "*innocence*," the "pre-reflective condition" of ordinary human reason, is susceptible to be seduced (G 4:405).

He continues to suggest that the source of the "seducer" is human nature, particularly desires and inclinations whose joint satisfaction is called happiness. The claim to happiness might bring about a powerful "counterweight" against the indispensable aspect of morality (the command of duty given by reason) (G 4:405). Here it is important to be clear that he does not mean that the voice of reason is simply superseded by the superior force of desires and inclinations, rather, he means that the seduction which results from desires and inclination provides unauthentic (fake) reason, but it appears as if it were legitimate reason and deceives the ordinary man to act against "the uncompromising demand of duty" (Alison 2011, 142-43).

Kant introduces the idea of *natural dialectics* to clarify the point that the issue to be considered is not a conflict between reason and sensibility rather competing reasons to act (he does not consider it as antinomy). As he describes it, this dialectics is: "a propensity to rationalize against those strict laws of duty and to cast doubt upon their validity, or at least upon their purity and strictness, and, where possible, to make them better suited to our wishes and inclinations, that is, to corrupt them at their basis and to destroy all their dignity something that even common practical reason cannot, in the end, call good (G 4:405)." More specifically, it is a dialectic of practical

reason in that empirically conditioned practical reason forms a deceptive illusion to improperly replace pure practical reason (Allison 2011, 143-44). This may bring danger to morality in practice and Kant puts the antidote not simply in philosophy or metaphysics of morals but in a "complete critique of our reason", and this critique begins in such a way that, as Kant asserts, ordinary human reason finds itself impelled or provoked on purely practical grounds to leave its original place and enter into practical philosophy "to obtain there information and distinct instruction regarding the source of its principle and the correct determination of this principle in comparison with maxims based on need and inclination (G 4:405)." According to Kant, this information is vital for ordinary human reason to get out of the difficulties of being deprived of its genuine ethical principle. Kant suggest that even common practical reason, "when it cultivates itself", knows that it is entangled into dialectic (G 4:405).

As we have seen above, the two concerns that makes common human reason need help from philosophy are "establishing the source and providing the correct determination of its moral principles" (Allison 2011, 144). More specifically, it shows that this principle is grounded in the concept of rational agent and specify its formula. These tasks are out of the territory of common human reason, and it is a task of philosophy, for Kant. From this, we can perceive that the solution to the difficulties facing common human reason lies in the *critique of our reason*. What gives a basis for this critique is Kant's account of the "plight of common human reason". To make clear what this critique is, it is the critique of an empirically conditioned practical reason when it attempts to affirm its authority over practical reason as such which includes pure practical reason as well. The purpose of this critique is to show that when empirically conditioned practical reason pretends in this way, it will overstep its own boundary by attempting to obscure the real difference

between it and pure practical reason. This also constitutes Kant's criticism of popular moral philosophy (Alison 2011, 144).

Chapter Two

2. The Background and the Formulation of the Categorical Imperative

2.1. Overview

Kant, in section two of the *Groundwor*k, begins with the criticism of popular morality and continues to the exposition of the supreme principle of morality. In section one of the *Groundwork*, he exposed the supreme principle of morality by analyzing ordinary views of morality. Thus, the fact that he has already accomplished this task in section one of the *Groundwork* gives rise to the question: why he continues the search and to establish the supreme principles of morality? The most important answer to this question, according to Hennery Allison, would be based on the nature of the supreme principle of morality. The supreme principle of morality has the characteristics of absolute necessity. This nature or status of the principle cannot be shown by claiming that it is implicitly found in common human reason which is Kant's objective in section one of the *Groundwork*. It is because to say this is tantamount to say that this principle is grounded in human nature which in turn prevent it from possessing the required status (its absolute necessity). Thus, in the following section, Kant aims at showing the supreme principle of morality has this characterstics. To accomplish this, he analyses the concept of a rational agent as such, more specifically, a concept of finite rational agent and he derives it from the analysis of this concept (Allison 2011, 149-150).

Kant makes, however, the situation more complicated. He does not stop in deriving the first formula of a categorical imperative, namely formula of universal law (FUL), rather he continues to derive additional formulas to what he names a single categorical imperative. But, what

purpose would be accomplished by the introduction of the additional formulas? According to Hennery Allison, for Kant, the additional formulas refer to different stages that in turn enable him to fully construct the concept of categorical imperative (Allison 2011, 150-51).

2.2. The Background of Categorical Imperative

2.2.1. Rational Agency and Imperatives

Kant begins his analysis of rational agent with the following passage: "Everything in nature works according to laws. Only a rational being has the capacity to act according to the representation of laws, that is, according to a principle, or a will. Since reason is required for the derivation of actions from laws, the will is nothing other than practical reason (G 4:412)."[8]

The above passage gives rise to different questions or different interpretations. The first question is about the nature of laws according to the representation of which rational agents act. In the first sentence, by "laws" Kant means laws of nature. From this, one may infer that Kant uses this meaning of law in the second sentence as well. If so, the essential feature of rational agents would be that they act according to the representation of laws of nature rather than merely according to these laws. This account of rational agency, however, does not help to understand Kant's argument that aims at formulating the moral law in its imperatival form (Allison 2011, 151-52).

One may also think that since the aim of Kant is the formulation of categorical imperative, the law that Kant has in his mind is moral law. This interpretation, however, is too narrow, as the further course of the argument shows. This is to mean that although his long run goal is to formulate categorical imperative (moral law), he uses the term law, at this stage, as *objective practical principle* which includes both instrumental and moral practical principles (Allison 2011, 152).

[8] Although this passage is differently interpreted by many commentators of the *Groundwork*, I am following Hennery E. Allison (2011).

The third attempt is to identify the law according to the representation of which rational agents act with maxims or subjective principles[9]. However, this interpretation should also be ruled out because, as the ensuing discussion shows, it creates exogenesis difficulties. That means, Kant associates maxims with finite or imperfectly rational beings with incentives and interests. But, the attribution under consideration, at this stage, applies to rational agents as such that includes both finite and infinite rational beings. Thus, perfectly rational beings do not act according to their representation of maxim because they have no maxims and they are not subject to sensuous influences or they do not act on interests (Allison 2011, 152-53).

Thus, according to Hennery E. Allison's reading, the framework or context within which Kant's analysis is undertaken obliges us to take *objective practical principles* as the most genuine candidate of the term "laws" according to the representation of which rational agents act (Allison 2011, 153).[10] By objective, Kant means that they are valid for all rational agents as such (G 4:413). Since these principles are valid for all rational beings, they can be applied both to holy or perfectly rational agents and imperfect rational beings. The only difference is that for perfect rational beings they appear as descriptive laws, whereas they appear as prescriptive laws as far as imperfect rational beings are concerned.[11] In the case of finite rational beings, both the instrumental and moral

[9] This reading is mainly supported by H. J. Paton (1947, 80-81).

[10] Timmermann (2007, 60) and Willaschek (2006, 125) affirm this reading.

[11] However, as Timmermann(2007, 62-63, fn27) affirms it, it is difficult even strange to think that the divine will can stand under laws that correspond to hypothetical imperatives, specifically to pragmatic imperatives. In the case of technical imperatives, as Allison (2011, 153, fn16) suggests, the difficulty would be minimized if God is considered as "the creator and providential governor of the universe". Because if God is considered as so, He has to have ends and chooses the best conceivable means to attain the end. But, the difficulty with regard to pragmatic imperative remains since God does not have inclination which in turn creates difficulty to attribute happiness to Him.

principles are connected to maxims, because they prescribe norms on the basis of which finite rational agents ought to elect their maxims and actions (Allison 2011, 153).

The above mentioned passage also gives rise to a question: what does it mean when it is claimed that rational beings act according to their representation of laws instead of just according to laws, as compared to everything in nature? To explain this point, we have to emphasize again the fact that Kant starts his analysis of rational agency from the belief that everything in nature obeys laws of nature. For instance, if we throw a stone upward, it will come down to the ground at a certain point, because of the law of gravity. But, the stone does not say to itself I ought to go up and come down when it arrives at its highest point since the law of gravity obliges this. In other words, the stone is not capable of reflecting on the law upon which it acts; or it is unable to announce to itself the law according to which it acts. In contrast, rational agents do not act just according to laws, but they are also capable of reflecting on the laws on which they act. This is because they are not just natural beings but also rational creatures. They are even capable of announcing to themselves the laws upon which they act (Korsgaard, 1998, xvi). Thus, the phrase "act according to one's representation of laws" refers to an agent who recognizes a law as normative to him/herself since it is recognized as normative or a law for other rational agents in similar circumstances. More specifically, to act according to one's representation of laws means acting according to recognized norms which is taken by Kant as equivalent to "acting on principles" (Allison 2011, 154).

In the above passage, Kant also claims that if one is governed by this sense of law, he is said to have a will. He defines will "as the capacity to act according to the representation of laws (G 4:412". In this case, he identifies will as practical reason (G 4:412). Allison maintains that Kant applies this identification to rational agents in general that includes both perfect and imperfect

rational agents. However, Kant also holds that reason determines the will. For Hennery E. Allison, when Kant claim that practical reason determines or fails to determine the will, he is talking about only in relation to finite rational agents. In other words, this attribution applies only to finite or imperfect rational beings. In order to make this point clear, Kant provides us with two kinds of wills determined by reason or two ways in which reason determines the will. The first kinds of will is the holly or perfectly rational will which is infallibly or invariably determined by reason. In this kind of will, there is no any competing force that dictates it to go out of the right track. We can find this kind of will in holly or perfectly rational agents (Allison 2011, 154). These beings necessarily act according to laws of reason. For them, actions which are recognized as objectively necessary are also subjectively necessary since there is no countervailing force that makes the agent to act opposite to those actions dictated by reason. On the other hand, in the second case, we can find a will such as human will in which reason fails to invariably determine it because it is subject to subjective inclinations or inceptives. Subjective inclinations lead the agent contrary to what is considered by reason as objectively necessary. As a result, for this will, actions which are identified as objectively necessary are not subjectively necessary, rather they are subjectively contingent. Consequently, this kinds of will is subject to constraint or necessitated by the objective laws of reason (G 4:412-14).

As mentioned above, the human will and every other finite will are categorized under the second group. Accordingly, they are subject to constraint or necessitation. This takes us to the idea of imperatives. Kant attempts to clarify the issue in the following passage: "The representation of an objective principle, in so far as it is necessitating for a will is called a command (of reason) and the formula of a command is called [an] *imperative* (G 4:413)." From this definition, we apprehend

that when an objective principle necessitates or compels a will, it is called the "command of reason", and the *formula*[12] of such a command is said to be an imperative (Paton 1947, 114). According to Kant, all imperatives are expressed by the word "ought" or can be expressed by the word "ought". "Ought" indicates the necessitation that hold between an objective principle and imperfectly rational beings. He also suggests that all imperatives command actions which are considered as good (both instrumentally and morally) on the basis of objective ground of reason or by objective principles valid for all rational agents (perfect and imperfect rational beings). Accordingly, the good differs from the "agreeable" which is determined by subjective causes or sensation rather than objective principles which are valid for all rational agents (G 4:413).

Perfectly rational beings necessarily act on objective principles which are imperative for imperfectly rational beings. They manifest the same kind of goodness manifested when imperfectly rational beings act based on imperatives. Thus, these objective principles are not imperatives for them; they are necessary but not necessitating (G 4:414).

After discussing the nature of imperatives, Kant introduces the distinction between hypothetical and categorical imperatives. However, the distinction between these imperatives is not made on the basis of grammatical or logical forms because a moral imperative can be expressed in a hypothetical form and a non-moral imperative can be expressed in a categorical form. It is rather made by the way they command. That means, what is important is how an imperative commands than what it commands (Allison 2011, 156). As Kant suggests "All imperatives command either hypothetically or categorically (G 4:414)." To command hypothetically is to command

[12]Although Kant uses the term formula in a number of ways which are context dependent, he uses in moral principles as a rule for solving problems. But, here, since Kant has not yet distinguished the categorical imperative from hypothetical imperatives, he uses, in the present context, the term formula as the verbal formulation of command; imperative is the formula of a command of reason in the sense of being its verbal expression; see Allison 2011, 155-56.

something under certain condition. It is a command to achieve a certain presupposed end. By contrast, to command categorically is to command unconditionally independent of any calculated end. The two ways in which imperatives command (i.e., hypothetically and categorically) are connected to two senses of the good, i.e., good as a means to something else and good in and of itself respectively. In other words, hypothetical imperatives command actions that are instrumentally good, whereas categorical imperatives command actions that are good in and of itself (G 4:414).

Moreover, Kant expands his twofold distinction between imperatives to three to take into consideration the two kinds of hypothetical imperatives (Timmermann 2007, 65). He makes a distinction between two kinds of hypothetical imperatives based on whether the end that finite rational beings pursue is either contingent or non-contingent. For Kant, the only end which is necessary and non-arbitrarily natural is happiness. Hennery Allison reminds us that this distinction reflects two unique features of Kant's conception of happiness as an end. First, unlike other ends, happiness is an end that all finite beings with sensuous as well as rational nature necessarily or naturally have. Second, unlike other ends, it is indeterminate in the sense that it is the satisfaction of the sum-total of an individual's inclination (Allison 2011, 157). So long as happiness is an end that all finite beings necessarily have, the imperative to will the means that is necessary to attain it might seem to be a categorical imperative. However, this kind of imperative is merely hypothetical since Kant defines happiness as the satisfaction of the sum of an agent's inclination. That means although all finite beings have the formal desire that inclinations are to be satisfied, they have different inclinations. Thus, they have different conception of happiness. Even they follow different means to attain it. Consequently, there is no single, universal principle that dictates the means necessary to attain happiness (Timmermann 2007, 68).

"In an effort to introduce a systematic ordering into his account of imperatives, Kant characterizes the various imperatives in terms of the functions of modality from his table of the logical functions of judgment in the first Critique (Allison 2011, 158)." Thus, Kant suggests that when the end pursued is arbitrary or contingent, the imperative commands *problematically*. When the end is one willed by any finite rational being or when the end is happiness, the imperative commands *assertorically*. The last imperative commands *apodictically* since it commands unconditionally without considering any end. The first two imperatives refer to the two kinds of hypothetical imperatives identified by Kant, whereas the last one refers to categorical imperatives. Kant also characterizes the first kinds of hypothetical imperative (imperatives that command *problematically*) as technical imperative or imperatives of skill, the second types of hypothetical imperative (imperative that commands *assertorically*) as pragmatic imperative and categorical imperatives as imperatives of morality (G 4:415-17). As Timmermann shows, while the first attribution (*problematical, apodictical and assertorical)* refers to the manner that the three types of imperatives command, the latter refers to their domain namely skill, prudence[13] and morality respectively (Timmermann 2007, 65).

Kant holds that the above three kinds of imperatives exercise different kinds of necessitation or constraint. In order to make their difference clear, he names them as rules of skill, counsels of prudence and commands (laws) of morality respectively (G 4:416).

The rules of skill might be more definite and clear, but they can be binding only so far as one wants to attain a particular end. The counsels of prudence are in some sense more uncertain, but they are more binding. They are said to be more uncertain in the sense that different individuals

[13] As Timmermann (2007, 67) affirms it, Kant uses the word prudence in relation to consciously maximizing one's happiness.

may use different counsels. It is difficult or perhaps impossible to be certain from where an individual can find his/her happiness and this highly depends on what the individual believes to be necessary to attain his/her happiness. However, for Kant, despite its uncertainty, the counsels of prudence is more binding than rules of skill because happiness is not an end which is simply abandoned or it is not an arbitrary end (G 4:416). Kant holds that both rules of skill and counsels of prudence are opposed to commands of morality, and only commands of morality are entitled to be laws.[14] H. J. Paton suggests that Kant here uses the word "command" to show that moral obligations do not depend on inclinations or they may even be opposed to them (Paton 1947, 116).

However, Kant, later, dismisses his tripartite classification of imperatives. In the "*First Introduction to the Critique of the Power of Judgement"*, he admits the mistake he made in the *Groundwork* in calling the imperatives of skill *problematic* imperatives. The reason for this is that the notion of commanding *problematically* leads to a contradiction in "terms" or self-contradictory since it is not the imperatives themselves problematic when they apply but it is the end they presuppose. He, thus, says that they have to be called technical imperatives. He further claims that although imperatives of prudence presuppose non contingent or actual ends that are subjectively necessary, they have to be included in the technical imperatives (Fi 20: 200). The reason for this is that prudence is nothing other than "the skill of a being to use for one's intention for free human beings and among these even the natural dispositions and inclinations in oneself" (Fi 20: 200). Thus, Kant seems to affirm a binary classification imperatives. In support of this idea, Kant discusses imperatives in *Metaphysics of Morals* by classifying them into two as categorical or unconditional and technical or conditional imperative (Mm 6:222).

[14] Kant at G 4:413 affirms that all imperatives are both laws and commands which is directly opposed to the present claim.

2.2.2. How are Imperatives Possible?

After discussing the different kinds of imperatives, Kant raises the question: "How are all these imperatives possible? (G 4:417)." This question is conceptual. That means, Kant wants to consider how to understand the necessitation or constraint imposed upon the will by imperatives and also how it is possible for finite rational beings to be motivated upon them as rational commands of their will (Allison 2011, 157).

In order to understand the argument that Kant uses to answer the above question, we must understand the distinction between analytic and synthetic propositions. The predicate of any analytic proposition is contained in the subject concept of it and it can be derived by mere analysis of the subject concept. Thus, to justify any analytic proposition, it is not necessary to go beyond the subject concept. By contrast, in any synthetic proposition, the predicate is not contained in the subject concept and hence it is impossible to derive it by mere analysis of the subject concept. Thus, to justify a synthetic proposition it is a must to go beyond the subject concept and search for another third term that enables us to attribute the predicate to the subject concept (Paton 1948, 28).

According to Kant, the possibility of imperatives of skill do not need any special explanation because all of them are analytic propositions. That means they are grounded in "Whoever wills the end also wills (in so far as reason has decisive influence on his actions) the indispensably necessary means to it that are within his power" (G 4:417). Kant seems to assume that, as H. J. Paton suggests, the end is always a result or an object, and for an end, it has one means which is considered as some possible action of the agent. If so, we can say that in the concept of willing an end (not just desiring or wishing it), there is contained the concept of willing the action which is the means to the end. Thus, the proposition that "to will the end is to will the means is a theoretical analytic proposition" (Paton 1947, 124).

Now the issue is how this theoretical analytic proposition becomes practical analytic proposition and hypothetical imperatives. As H. J. Paton suggests, Kant himself is sketchy regarding this issue. According to Paton, to fill this gap, we have to remember the above objective principle of practical reason. Although this proposition is still analytic, it appears imperative for us since reason does not have a decisive influence although it exists in us. Thus, it can be given the form "if any rational agent wills the end, he ought to will the means" (Paton 1947, 124).

Moreover, Kant notes that to know the means of the proposed end, we have to use synthetic propositions. We are required to know the cause of a certain desired effect and it is impossible to know what the cause of a certain effect through mere analysis of the effect itself. These synthetic propositions are theoretical. However, the imperative is still analytic as far as willing is concerned. In other words, once we have this theoretical knowledge, the principle that governs rational action and imperatives of skill is the principles of: "Whoever wills the end also wills (in so far as reason has decisive influence on his actions) the indispensably necessary means to it that are within his power" (G 4:417). This proposition is analytic and it continues to be so in action as well although many specific synthetic propositions are required in order to determine the specific cause of the specific end proposed at a particular situation (Paton 1947, 125).

When we come to imperatives of prudence, it is not simple to justify them. Kant encounters a special difficulty as far as imperatives of prudence is concerned. Even though happiness is an end that every human being naturally seeks, its concept is vague and indeterminate. That means, different people attains happiness through different means. So, it is impossible to identify a means which is actually necessary for all human beings to be happy. However, apart from this difficulty, pragmatic imperative can be justified like technical imperatives (G 4:418-19).

The above kind of justification is not applicable to categorical imperative. A categorical imperative is unconditional and does not refer to an end that the agent seeks to attain (G 4:419). In order to justify a categorical imperative, one has to show that "a fully rational agent would necessarily act in a certain way, not if he happens to want something else, but simply and solely as a rational agent" (Paton 1947, 127). This proposition is not analytic because the predicate of this proposition is not contained in the subject concept 'rational agent' and it cannot be derived by mere analysis of the subject concept. Thus, this proposition is synthetic. Moreover, it is an assertion of what any rational agent ought to do. As a result, it cannot be justified by appealing to experience. Hence, the categorical imperative is not only synthetic but also a priori. And it is very difficult to justify it. Kant postpones this task to section III of the *groundwork* (Paton 1947, 127).

2.3. The Formulation of the Categorical Imperative

2.3.1. The Formula of Universal Law (FUL)

As noted earlier, Kant announces that he postpone the question of the possibility of categorical imperative to section III of the *Groundwork*. However, although the justification of this standard as the objective principle of rational agents cannot be derived from its very concept, it is quite possible to derive its content or what it enjoins from its concept (Timmermann 2007, 73).

Since hypothetical imperatives are conditional, they command based on the end that the agent wants to fulfill. The necessitation or constraint imposed by these imperatives are means to the agent's presupposed end. Thus, hypothetical imperatives as such do not command anything; that is, such imperatives command the agent so long as a particular end is provided. Thus, the agent needs to wait until the end is provided (G 4:421).

On the other hand, categorical imperative is unconditional. The constraint or necessitation imposed upon the will is independent of any end presupposed by the agent. Otherwise expressed,

it just commands us to act (to adopt maxims[15]) in conformity with objective principle or "practical laws" (G 4:421). Laws are equally valid for all rational agents as such and they are universally and unconditionally valid constraints on action. Thus, the conformity of maxims to universal law as such follows from the mere concept of categorical imperative (G 4:421fn). However, the maxim of finite rational beings can conform to universal laws, according to Kant, if and only if such agents can also will them as universal laws. What is important is the compatibility between the material maxim (actual maxim) and the thought of the same maxim as universal laws. From this idea, one can understand that a categorical imperative requires that an agent adopts only the maxim that he/she can at the same time will as a universal law. This can be expressed in the formula: "*Act only according to that maxim through which you can at the same time will that it becomes a universal law"* (G 4:421). The term "formula", here, according to Hennery Allison, refers to a rule for solving a problem. The problem is to make clear how the law can be applied to imperfect (finite) rational beings, and the answer to this problem is given in the form of a categorical imperative (Allison 2011, 174).

Kant holds that the above formulation is the one canonical and general formulation of the categorical imperative. For him, there is only a single categorical imperative and it is the above principle (FUL). Other unconditional imperatives are either variants of this principle or individual 'categorical imperatives', i.e., particular applications of this principle, like an unconditional command not to kill. The unique status of the formula of universal law is asserted when Kant summarizes the three variants at G 4:436.The formula of laws of nature with its four illustrations, which immediately follows the present formulation, is the first of three variants. The fact that there is

[15] Kant assumes that every action, as an intentional deed of a rational agent, involves the adoption of a maxim.

only one supreme law is again a consequence of its nature as an unconditioned practical law (Timmermann 2007, 76-7).

The above test, however, is abstract and Kant thinks that it may be difficult to apply it. He realizes that it is easier to apply the test to a maxim if we consider it not as a normative law but as a law of nature, a universal rule against which it is causally impossible for everyone to act (Wood 1999, 80). In the next section, I shall discuss the formula of law nature and how Kant makes the relation between this formula and FUL.

2.3.2. The Formula of Law of Nature (FLN)

In the *Groundwork of the Metaphysics of Morals*, a few lines after stating FUL, Kant suggests that:

> Since the universality of law in accordance with which effects take place constitutes what is properly called nature in the most general sense (as regards its form) - that is, the existence of things in so far as it is determined in accordance with universal laws - the universal imperative of duty can also go as follows: *act as if the maxim of your action were to become by your will a universal law of nature.* (G 4:421)

Kant writes the above passage based on the distinction that he made in the C*ritique of Pure Reason*. He conceives of nature formally and materially. When nature is conceived formally, it is the existence of things according to universal laws. What Kant implicitly wants to claim is that this concept of nature is used to show the conformity to universal law assumed in the categorical imperative. He makes the transition from FUL to FLN (Allison 2011, 177). According to Hennery Allison, this immediately gives rise to the question: "How can the idea of conformity to laws of nature represent the idea of conformity to laws of a completely different type, specifically, laws of freedom? (Allison 2011, 177)"

It is possible to address this question in the following way. In FUL, Kant informs us that the maxims on which we act ought to conform to universal law as such. He then continues to tell

us that the *form of nature* is created by the universality of laws. That means, in nature, every event takes place according to universal law, particularly the law of cause and effect. From this, we can understand that nature and moral action have the same form although the law of freedom and the law of nature are not the same. This indicates that there is analogy between universal law of freedom and universal laws of nature. It is through this analogy that Kant assumes that the universal law imposed by the categorical imperative is represented by the law of nature and goes from FUL to FLN (Paton 1947, 157-58).

In *Critique of Practical Reason*, Kant affirms that since FLN contains the key idea of universality, it is characterized as a "typic" of FUL. By this, he means a model or a symbolic representation by which we apply FUL to a particular maxim in moral deliberation (cpr 5:66-70).

To use this formula, Kant tells us to imagine ourselves as a creator of a world to which we are a part. In the *Critique Practical Reason*, he makes it more clear by stating that "ask yourself whether, if the action you propose were to take place by a law of the system of nature of which you were yourself a part, you could regard it as possible by your own will (cpr 5:69)." The most important point to be clarified here is, whether the proposed course of action is taken as a law of nature in which the agent is a member. The idea of being the law of nature expresses the idea of universality required by the categorical imperative (Allison 2011, 178).

However, as Hennery Allison suggests, there is ambiguity with regard to the precise function assigned to the FLN by Kant. Some of Kant's claims suggest that the FLN is intended to function as a self-standing source of duties or as determining duties.[16] For instance, the claim that "all imperatives of duty, that is, all imperatives specifying particular duties, can be derived from

[16] Jens Timmermann (2007) affirms this function of FLN.

FLN as their principle" strengthens this idea. On the other hand, Kant's discussion of the four examples in relation to the four kinds of duties, in the *Groundwork*, suggests that the function of FLN is to rule out maxims that violate each of the four duty types to be discussed next[17] (Allison 2011, 179).

2.3.3. The Application of FLN to the Four Examples of Duty

After stating the FLN, Kant enumerates four types of duties: perfect or strict or narrow duties, imperfect or wide duties, duties to ourselves and duties to other human beings. He gives one example or illustration for each to show that this formula can be applied to all kinds of duties mentioned above (G 4:421)[18].

He defines perfect duty as a duty that does not admit of any exception in favor of any inclination, whereas imperfect duty is one that admits of some exceptions in favor of inclination. Kant makes the distinction between perfect and imperfect duty in another way. In the case of the former, one is dictated to do a definite act, whereas in the case of the latter, he/she is dictated to follow merely a maxim. Or imperfect duties are duties to set ends. One is expected to fulfill these ends, but it is up to him/her what and how much he/she does to fulfill them. The particular action he/she uses to fulfill those ends is meritorious and any failure to act in a particular way represents only a lack of merit rather than a piece of wrong doings. Actions are wrong if and only if they fail to fulfill the required end by presupposing ends which are contrary to the required one (Mm 6:227).

---

[17] Hennery Allison (2011) supports this view.

[18] As he indicates in a note attached to this passage, in the *Groundwork*, Kant proposes to reserve a systematic taxonomy of duties for a future *Metaphysics of Morals*. Thus, the present classification is considered only discretionary; that is, he makes the present division as a convenient device for ordering his examples (G 4:421fn).

With regard to imperfect or wide duties, there is a *latitude* that helps one to determine which action and how much he/she performs it in order to fulfill the required end. It is within this latitude that inclinations are entitled to play a crucial role in determining what the agent performs (Mm 6:390). For instance, the duty of benevolence to a neighborhood might be limited by similar duty to one's parents. Thus, it is an example of an imperfect duty (G 4:423). On the other hand, the duties of paying a debt cannot be limited by another. Because one cannot refuse to pay his debt whatever inclination he has. Thus, payment of debt is an example of perfect duty (G 4:422).

After enumerating the type of duties, Kant suggests that maxims that violate perfect duties cannot be thought or conceived without contradiction as a universal law of nature. However, maxims that violate imperfect duties can be conceived without contradiction, but they cannot be willed without contradiction (G 4:424). They are considered, in literatures, as "contradiction in conception" and "contradiction in will"[19] tests. I will discuss them in the following pages in relation to the examples they are expected to cover.

The first illustration is the problem of committing suicide. This example is given under the heading of perfect duty. In the *Groundwork*, Kant writes that: "someone feels sick of life because of a series of troubles that has grown to the point of despair, but is still so far in possession of his reason that he can ask himself whether it would not be contrary to his duty to himself to take his own life (G 4:422)." Kant points out that, the maxim of the agent who wants to commit suicide is that: "from self-love I make it my principle to shorten my life when its longer duration threatens more troubles than it promises agreeableness (G 4:422)." According to Kant, now the agent asks whether his maxim can be a universal laws of nature. To check this, we have to see the universalized form of the above maxim. As Allen Wood formulates it, its universalized form is: "it is a

[19] These terminologies are devised by Onora O'Neill; for this, see (Wood 1999, 84).

universal law of nature that from self-love, all make it their principle to shorten life when by a longer duration it threatens more ill than agreeableness" (Wood 1999, 85). It is important to understand that Kant's argument that this universalized form of the maxim cannot be conceived as a law of nature is not based on the claim that this maxim is directly contradictory to itself. Because, as Hennery Allison reminds us, "a world in which there was a law of nature specifying that everyone acts on this maxim would be significantly depopulated, but there is no contradiction in that" (Allison 2011, 184). Rather, the law of nature that could be established by universalization of the above maxim is teleological. So, the contradiction that Kant has in his mind is a teleological contradiction. According to Christine M. Korsgaard, such kind of contradiction occurs when something is used in a way which is inconsistent to its natural purpose or used in contrary to what its natural purpose needs (Korsgaard 1996, 87). Thus, the contradiction occurs on the basis of the alleged fact about the natural purpose of the feeling of self-love; that is, the natural function of this feeling is to further life (G 4:422). Here Kant argues that:

> A nature whose law it was to destroy life through the same sensation whose proper function is to impel the furtherance of life would contradict itself, and thus could not subsist as nature; hence that maxim could not possibly hold as a universal law of nature and, consequently, entirely contradicts the supreme principle of all duty. (G 4:422)

The second and most widely discussed examples of Kant is false promise. He puts the issue in the following way:

> Another finds himself urged by need to borrow many. He well knows that he will not be able to repay it but sees also that nothing will be lent him unless he promises firmly to repay it within a determined time. He would like to make such a promise, but he still has enough conscience to ask himself: is it not forbidden and contrary to duty to help oneself out of need in such a way? Supposing that he still decided to do so, his maxim of action would go as follows: when I believe myself to be in need of money I shall borrow money and promise to repay it, even though I know that this will never happen. (G 4:422)

The next point is that this principle of benefiting oneself might be consistent to the whole future welfare of the agent. But, he is not certain whether it is right. Therefore, the agent turns this

principle of self-love to a universal law and he applies the FLN to answer the question of whether he can will the maxim, without contradiction, as a universal law of nature. Unfortunately, the maxim in question cannot be consistent with itself and even contradicts itself. Thus, it is unable to be entitled as a universal law of nature (G 4:422). Kant expresses this point more elaborately in the following way: "for, the universality of a law that everyone, when he believes himself in need, could promise whatever he pleases with the intention of not keep it would make the promise and the end one might have in it itself impossible, since no one would believe what was promised him but would laugh at all such expressions are vain presence (G 4:422)."

The above example poses a question with regard to the nature of the contradiction that occurs when the above impermissible maxim is taken as a universal law of nature. In the literature, it is possible to find two interpretations of the kind of contradictions that Kant has in mind as far as the above example is concerned. The first interpretation claims that the kind of contradiction, contradiction in conception, involved in the above example is "logical contradiction[20]". This contradiction occurs when one wills "a" and not "a" at the same time. Thus, since the agent is willing that there would both promise and no promise, the contradiction that Kant assumes to appear in the above example is logical contradiction, according to this interpretation. The agent is willing that there is promise because he is claiming to earn the needed money in the name of promise. At the same time, he is willing that there would no promise because the agent is making a promise without the intention of making it, and such kind of promise would not count as a promise (Allison 2011, 186).

[20] See Korsgaard (1996, 81-82) and Wood (1999, 87-89).

The second and most recent interpretation of contradiction that a false promising example entails is "practical contradiction[21]". According to this interpretation, the contradiction is practical because it takes place between the end that is attained by making a false promise and the state of affairs that would happen if false promising were a universal law of nature. Expressed otherwise, the end that would be attained by false promising cannot be attained since a promise, the means to attain the end, will not be believed. This is a genuine practical contradiction that occurs when the agent wills both the end and the state of affairs that make it impossible at the same time (Allison 2011, 186).

Kant gives his third example in relation to a failure to develop one's talents or capacities. He argues:

> A third finds in himself a talent that by means of some cultivation could make him a human being useful for all sorts of purposes. However, he finds himself in comfortable circumstances and prefers to give himself up to pleasure than to trouble himself with enlarging and improving his fortunate natural predispositions. But he still asks himself whether his maxim of neglecting his natural gifts, besides being consistent with his propensity to amusement, is also consistent with what one calls duty. (G 4:423)

Kant takes cultivating one's talents as an imperfect duty to oneself. He call the maxim that advocates their neglect as "a maxim of slothfulness". He considers the cultivation of one's talents as an imperfect duty since the maxim that advocates the neglect of such talents will lead to a contradiction in willing rather than in conception (Allison 2011, 187). The contradiction that appears in this maxim is in the will rather than in the conception is made clear by Kant in such a way that: "a nature could indeed always subsist with such a universal law, although … the human being should let his talents rust and be concerned with devoting his life merely to idleness, amusement, procreation-in a word, to enjoyment" (G 4:423). The question here is why the agent cannot will the maxim of slothfulness as universal law of nature? Kant answers that the agent does not will this

[21] See Korsgaard (1996, 92-94).

maxim as a universal law of nature because "as a rational being he necessarily wills that all the capacities in him be developed, since they serve him and are given to him for all sorts of possible purposes (G 4:423)".

Like the previous example, there is a debate between the teleological and practical contradiction interpretations with regard to the nature of contradiction that Kant holds in his mind in this example. The teleological contradiction interpretation is based on Kant's characterization of the talents or capacities that the slothfully person wants to ignore as "fortunate natural predispositions" or "gifts of nature". By considering nature as the source of these talents or capacities, it is possible to expect a teleological contradiction which is analogues to a contradiction that Kant appeals in the example of suicide. The only difference between these teleological contradictions is that the maxim of slothfulness is ignoring rather than acting directly against something which is a gift of nature since it is an infringement of an imperfect duty. This again explains the reason why it can be conceived, although cannot be willed, without contradiction (Allison 2011, 188).

The practical contradiction interpretation is dominant as compared to teleological contradiction interpretation. This reading or interpretation, however, has its own problem because it is reduced to prudential matters. If so, it involves a violation of hypothetical rather than categorical imperatives. Thus, from prudential point of view, as suggested by Hennery Allison, the argument would be that "a maxim of neglecting to develop one's capacities in order to maximize the agreeableness of one's present condition might be seen as a case of willing an end (an agreeable life) without willing an essential means to this end, namely capacities sufficiently developed to ensure its maintenance (Allison 2011, 188-89)".

Kant's fourth example is given in relation to non-beneficence. In this example, he wants us to imagine a person who is indifferent to interfere in the affairs of his fellow human beings. He

puts the issue as follows: “for whom things are going well while he sees that others (whom he could very well help) have to contend with great hardships, thinks: what is it to me? Let each be as happy as heaven wills or as he can make himself; I shall take nothing from him nor even envy him; only I do not care to contribute anything to his welfare or to his assistance in need! (G 4:423)” Kant indicates that there is no contradiction in conceiving such a way of thinking as a universal law of nature since human beings can still exist (G 4:423). However, he suggests that the agent could not will this principle as a universal law of nature, because “a will that decided this would conflict with itself, since many cases could occur in which one would need the love and sympathy of others and in which, by such a law of nature arisen from his own will, he would rob himself of all hope of the assistance he wishes for himself (G 4:423).”

This example depends on the assumption that there are ends that a finite rational being cannot abandon since they essentially belong to the will. If not, it would be open to an agent to dismiss the assistance which comes from other agents and this in turn enables the agent to be absolutely independent. For instance, hanger, at least under some circumstance, cannot be abandoned by a finite rational agent. Thus, if this attitude is adopted as a universal laws of nature, it involves a practical contradiction since such an agent is willing at the same time both the no one would help those who are in need and others must be willing to help him when he is in need (Timmermann 2007, 85).

2.3.4. The Formula of Humanity (FH)

Kant moves to the FH by enriching his previous idea of rational agency (Allison 2011, 204). As noted before, he holds that rational agents are beings with the capacity to act according to their representation of laws. He equalizes this capacity with the possession of a will or practical reason (cf. G 4:412). He re-expresses this idea in slightly different words by saying: “The will is

thought as a capacity to determine itself to acting in conformity with the *representation of certain laws*. … . Now, what serves the will as the objective ground of its self-determination is an end… (G 4:427)". The latter sentence brings about the second fundamental feature of all actions. That means, any rational action, in addition to having a principle, sets an end before itself (G 4:428).

According to Kant, ends might be either subjective or objective. Subjective ends are based on inclination. It is obvious that different agents have different inclinations. Thus, these ends are not based on reasons which is valid for all rational agents. Kant also identifies subjective ends with incentive which is the subjective ground of desire or inclination. These ends have a relative value and conditioned since their value is the function of the agents desire. So, these ends are only ground for hypothetical imperatives (G 4:428).

On the other hand, objective ends are based on pure reason. They are not adopted on the basis of inclination or personal preference but rather based on reasons that are valid for all rational agents. They have also unconditional and absolute value. This implies that these ends cannot result from mere human action since ends that result from mere human action do not have absolute or unconditional worth. Thus, they have to be an already existent end. Thus, these ends could be the ground of categorical imperative. They can also be expressed as ends in themselves, meaning they are not ends just in relation to a particular agent (G 4:428).

After having made clear the nature of ends that be able to ground the categorical imperative, Kant suggests that: "Now I say that the human being and in general every rational being *exists* as an end in itself, not *merely as a means* to be used by this or that will at its discretion; instead he must in all his actions, whether directed to himself or also to other rational beings, always be regarded *at the same time as an end* (G 4:428)." As we understand from the above passage, it is only rational beings or persons that can be ends in themselves. Since only persons have absolute

or unconditional value, it is not right to use them merely as means to an end its value is only relative. If there is no such ends, there would be no unconditioned good, supreme principle of actions or categorical imperative for human beings. Thus, the peculiar end which is required for the categorical imperative to be possible is persons or rational beings (G 4:429). To derive the new formula of the categorical imperative, namely the formula of humanity, Kant uses the idea that rational agents are ends in themselves. Thus, the FH looks like the following: "*So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means* (G 4:429)."

Here, "humanity" refers to the rational capacities of human beings. Both in ourselves and in others, we must respect and further the element or activity that enables us to be a human being (Timmermann 2007, 96).

2.3.5. The Formula of Autonomy (FA)

The formula of autonomy is formulated on the basis of the idea that a rational will makes or gives itself the law that it obeys (Paton 1947, 180). The formula is expressed as follows: "*So act that your will can regard itself at the same time as making universal law through its maxim* (G 4:437)."

At first glance, this formula seems a mere repetition of FUL. Because, in commanding us act only on a maxim through which we can at the same time will that it should become a universal law, FUL is suggesting that a rational will is making the law that it obeys. But, in FUL, the emphasis is given to the objectivity of the moral law and the necessitation imposed on the imperfect rational will. The FA makes explicit what is an implicit in FUL. Indeed, the FA does not deny the necessitation since it is a formula of duty. Nevertheless, what it makes clear is that the necessitation is exercised by the rational will itself. The will makes the law to which it is subject. It is subject to

the law because it makes the law. This formula is the most important formulation of the supreme principle of morality because it leads to the idea of freedom, for Kant (Paton 1947, 180).

Moreover, unlike the preceding formulas, the FA is not directly derived from the concept of rational being or rational will, but rather it combines the features of the categorical imperative discussed so far. This is the only thing which is clear. However, the detail of the argument that leads to the notion of autonomy are less clear since Kant tries to arrive at the FA in many ways that he does not clearly differentiate one another (Timmermann 2007, 102).

According to Paton, the FA can be derived from the FH in the following way: the setting of ends before oneself is the essential manifestation of freedom. One might be forced to do a certain action which can be directed as a means to some end. However, he/she cannot be forced or compelled by any other to make something his/her end. If he/she makes something his/her end, he/she does so by his/her own free will. If FH enjoins, indeed it does, the adoption of certain ends, in particular the treating of all rational beings as ends in themselves, then the necessitation present in duty must come from one's own free and rational will. That means, the agent must be the source and author of the law that he/she obeys. Therefore, FH, in commanding the pursuit of ends, implicitly asserts the autonomy of the will in making or giving itself the laws that it ought to obey (Paton 1947, 181).

According to Jens Timmermann, FA can also be constructed by combining the FUL and FH in the following way. Any act of volition has two elements: formal and material. The formal element is the law that objectively governs the action, whereas the material element is the end that the agent (the subject) freely adopts. In the case moral action, the rule or the maxim can be willed as a universal law as it is shown in the FUL. The FH also shows that there are ends in themselves as subjects contain all other ends, and every rational being is such an end. Thus, by combining the

objective and subjective aspect of the two formulas (FUL and FH), Kant leads to the idea that every rational being is involved in making universal laws (Timmermann 2007, 102-03).

There is another argument that shows how Kant derives the FA from the essence of categorical imperative. As we have seen so far, the moral imperative is categorical and unconditional, and thus it avoids any interest. The preceding formulas implicitly show this; they implicitly exclude interests (pathological interest). Now, the FA explicitly avoids any interest. That means, in suggesting that a moral will is autonomous [or makes the law that it obeys], it is already suggesting that it is not determined by any interest since a will determined by interest, desire and inclination is heteronymous. If will were determined by interest, desire or inclination, it would be subject to the law that it does not make, and ultimately to natural law (G 4:432-33). For Kant, all philosophers who want to determine moral obligation by any kind of interest make the categorical imperative inconceivable and dismiss morality altogether. These philosophers advocate the doctrine of heteronomy rather than the doctrine of autonomy; the will is bound only by laws originated from outside force rather than from the will itself. Thus, this kind of theory does not give rise to categorical imperative, rather it gives rise to hypothetical and non-moral imperative (G 4:433, 441).

Therefore, "we can say if there is a categorical imperative, the moral will which obeys it must not be determined by interest, and therefore must itself make the universal laws which it is unconditionally bound to obey (Paton 1947, 182)."

2.3.5.1. The Realm[22] of Ends

Kant suggests that the notion of autonomy further leads to another closely connected world, i.e., the realm of ends. He expresses this issue in detail in the following way: "The concept of every rational being as one who must regard himself as giving universal law through all the maxims of his will, so as to appraise himself and his actions from this point of view, leads to a very fruitful concept dependent upon it, namely that of a realm of ends (G 4:433)." As we understand from this passage, Kant considers the realm of ends as an idea by which one can construct for himself/herself the conception of autonomous agency. Otherwise expressed, when one considers him/herself as an autonomous agent, he is at the same time think of him/herself as a law-giving member of the realm of ends. To understand this idea more properly, it is essential to know what Kant means by "a realm of ends" (Allison 2011, 242). Kant defines realm as "a systematic union of different rational beings under common laws (G 4:433)". The term "systematic" here has a teleological connotation that indicates the various members live harmoniously and they also support one another. When we apply this conception to ends, the realm of ends is one in which the members are ends in themselves (Allison 2011, 242).

Nevertheless, the situation becomes more complex since the realm does not only include rational beings as ends in themselves, but also the personal or private end of these rational beings. This assertion seems to be inconsistent with what Kant says at the same sentence, namely that we reach the idea of the realm of ends by abstracting from the personal differences among the rational

[22] Although there is a general consensus among the English translators that the word "realm" is more appropriate than "kingdom" to the German word "*Reich*", there is a disagreement on the question that whether in Kant's use of the term it should be translated as "realm" or "kingdom". In this thesis, I use realm because of Kant's definition of *Reich*, to be discussed below; for detail discussion with regard to the issue, see (Paton 1947, 187-88).

beings and the content of their private ends.[23] However, according to H. J. Paton, what Kant means by this might be that when one acts as a law making member of the realm of ends, his/her action cannot be determined by personal differences or private ends. The realm of ends regard private ends only as long as they are compatible with universal laws. The system of ends, governed by self-imposed and objective laws, is a framework in which both the private end of ourselves and others are realized (Paton 1947, 187).

Kant also insists that the realm of ends is possible according to "the above principles (G 4:433)". By this, as Allison suggests, Kant may mean according to the preceding formulas of categorical imperative. However, in the next paragraph, he specifically connects this ends with the FH in saying that the realm of ends arises (Allison 2011, 243) "when everyone treat himself and all others never merely as means but always at the same time as ends in themselves (G 4:433)". The other important point is that Kant admits that the realm of ends is only an ideal, but his claim that it is possible in accordance with the categorical principle, as we have seen above and his later assertion that it could be come if everyone carefully obeys this principle "indicates that he viewed it as an ideal that could be realized (or at least approached asymptotically)" (Allison 2011, 243).

Moreover, there is another complication that results from Kant's ambiguous conception of the realm of ends. To express the point more clearly, although Kant considers the realm of ends as an ideal, he also considers rational agents as members of this realm. In fact, according to Kant, the realm of ends has ordinary members and the supreme head. Since both of them are ends in themselves, they make universal law. Their difference is that ordinary members do not merely make

[23] It is, here, better to quote Kant's idea directly to clearly see the inconsistency: "…if we abstract from the personal differences of rational beings as well as from all the content of their private ends we shall be able to think of a whole of all ends in systematic connection (a whole both of rational beings as ends in themselves and of the ends of his own that each may set himself), that is, a realm of ends…(G 4:433)"

laws, but they are also subject to the laws that they themselves make (G 4:433). Kant suggests that the supreme head is not subject to the will of any other which implies that ordinary members are subject to the will of others (cf. G 4:433). If this is Kant's real assertion, it is inconsistent since, for him, all autonomous agents are not subject to the will of another. Thus, this cannot be taken as a characteristics that differentiates the supreme head from the other. For Allison, it is not Kant's real assertion (Kant, here, misstates his true position). Allison argues that Kant's real idea, as it becomes obvious from the idea that follows, is that the supreme head is free from any desire and subjective inclination. So, the head is not subject to the moral law, whereas the finite rational beings that constitute the ordinary member are subject to the law (Allison 2011, 244)

Although the members, unlike the supreme head, are subject to the law, they are however subject to the law that they make. The realm of ends becomes possible only when its members are autonomous or it is possible only through freedom of will. Because they are autonomous (as law making member of realm of ends) members of such a realm have dignity; that is, they have an intrinsic, unconditioned, absolute, incomparable worth or value. At this point, Kant turns from the description of abstract values to the illustration of values (G 4:434).

Kant further argues that "in the realm of ends everything has either a *price* or a *dignity*." Price is a value of a thing in exchange. In relation to this, a thing has a price if it has an equivalent. But, it has a dignity or worthiness if it does not have such an equivalent or substitute. In the realm of ends, ends set by rational beings has a price, whereas ends humanity as ends in themselves has dignity (G 4:434).

2.3.5.2. The Formula of the Realm of Ends (FRE)

Having discussed the realm of ends, let me proceed to spell out the formula of the realm of ends. This formula, along with the FA, is the most vital formulation of Kant's categorical imperative. This formula can be expressed as: "*Act in accordance with the maxims of a member giving universal laws for a merely possible realm of ends (*G 4:439; cf. 4:438, 434)." Kant introduces this formula in combination with the FA and, in some passages, Kant even seems to consider FRE as expressions of FA (G 4:436). Moreover, FRE can be derived by combining the ideas of previously mentioned formulas of categorical imperative. Thus, this enables the mentioned formula to be more adequate in expressing the spirit of the supreme principles of morality (Wood 1999, 167).

2.3.6. The Connection between the Formulas of Categorical Imperative

Kant says that: "The aforementioned three ways of representing the principle of morality are at bottom only so many formulas of the very same law, and one of them of itself unites the other two in it (G 4:436)." Although this sentence is single, it contains two claims that require a separate treatment. The first claim is that these formulas are the expressions of the same law or principle. As Hennery Allison suggests, the law represented by the three formulas (FLN, FH and FA/FRE) are the single categorical imperative that first stated at G 4:421 and till now has been referred as FUL (Allison 2011, 246).[24] We saw before that Kant characterized categorical imper-

[24] The categorical imperative (FUL) is the fundamental principle of morality (moral law) which is expressed or formulated in its imperatival mood. However, so long as Kant considers the fundamental principles of morality in its imperative mood (the categorical imperative) a formula (FUL), this implies that the three formulas are formulas of a formula. And this seems incoherent. Hennery Allison try to avoid this incoherence by considering FLN, FH, and FA/FRE as first order formulas

ative as single. Thus, the above sentence strengthens Kant's claim that there is only one fundamental principle of morality. Kant warns his readers not to take the various formulas as distinct or independent ethical principles, rather they are the reformulations or variant formulations of the single categorical imperative (cf. Timmermann 2007, 110). In *Lecture on Ethics*, which supports the above claim, in condemning the plurality of other ethical theories, Kant says that "Where there are already many principles, there are certainly none, for there can only be one true principle (LE 27:266)".

The second claim contained in the single sentence mentioned above is that one of the three reformulations unites the other two in itself. It is in this context that Kant talks about the mutual relation between the various formulas of the categorical imperative. In order to show which formula contains within itself the other two, and thereby to explain the mutual relation between them, Kant claims that every maxim has a form, a matter (end) and a complete determination. He suggests that every maxim has the form of generality in the sense that it is a general principle or policy of action by which it is done. From this, Kant concludes that "in this respect the formula for moral imperative is expressed thus: that maxims must be chosen as if they were to hold as universal laws of nature (G 4:436)". By the law of nature, he equates it with FLN which specifies the formal condition that a maxim must fulfill that in turn enables to conform to the moral principle. In the same way, he argues that every maxim has a matter or an end to which the action is done (G 4:400). From this, he derives the FH in such a way that "a rational being, as an end by its nature and hence as an end in itself, must in every maxim serve as the limiting condition of all merely relative and arbitrary ends (G 4:436)". Thus, FH identifies the moral principle by

directly applied to maxims and FUL as second order formula which underlies and expressed by these first order formulas; for this, see Allison 2011, 253.54; see also Timmermann 2007, 104, fn108.

objective end and this end is the existent end of humanity as an end in itself. This end serves as a motive for a will that follows a categorical imperative. However, "complete determination" does not apply to an individual maxim, rather it is a necessary condition of an entire system of moral legislation that governs the conduct of a rational agent (Wood 1999, 184). The FA formulates the moral law in terms of this system. This becomes obvious when FA is expressed in the guise of FRE, because this formula explains that the law must unite all rational beings as ends in themselves into a harmonious organic system. In this way, FA/FRE not only goes beyond FLN and FH, considered individually, but also unites them. This also implies the interdependence and collective completeness of the three formulas (G 4:437).

Kant indicates that there is a progression from one formula to another and the three formulas complement one another in their practical application (G 4:436).

The three formulas refer to complementary ways when moral principle applied to maxims. FLN refers to the universal validity of each maxim; FH directs us to the many rational beings who must be treated as ends in themselves; and FA in the guise of FRE presents these ends as an organic system or a harmonious commonwealth (cf. Wood 1999, 185).

Kant argues that there is a progression from FLN to FH and then to FA/FRE. The FLN, as the legislative form of a maxim, leads to the search for the objective end which represent this legislative form. This end is found in rational beings as ends in themselves. Thus, FLN progresses to FH. The combination of FLN and FH provides FA/FRE in the sense that the worth of rational nature which grounds FH and conceived as the idea of a rational will is considered as the author of laws presupposed in FLN (Wood 1999, 185-86).

Finally, Kant recommends the strict method of FUL in practical purpose or for moral appraisal, but the three formulas (FLN, FH and FA/FRE) come to secure acceptance for the moral law or to bring the FUL to intuition through analogy (Timmermann 2007, 112).

Chapter Three

3. The Justification of the Categorical Imperative

3.1. Overview

As we have seen so far, in the first two sections of the *Groundwork*, Kant describes the procedure he used as analytic or regressive (Allison 2011, 273). In section one of the *Groundwork*, he starts from analysis of ordinary moral beliefs. Through this analysis, Kant arrived at a version of the categorical imperative, and he argued that the condition for moral action, at least for finite rational beings, is obedience to a categorical imperative. In section two of the *Groundwork*, he formulates the categorical imperative in at least three ways and he finished the search for the supreme principle of morality (Paton 1947, 199). Thus, at the end of section two of the *Groundwork*, in formulating the supreme principle of morality, Kant completed the first of the two tasks that he wanted to accomplish in the *Groundwork*, i.e., to search for and formulate the supreme principle of morality as it is found in common human reason (Allison 2011, 273).

Section three of the *Groundwork* is devoted to the second task that Kant wanted to accomplish; that is, to establish the objective validity of the supreme principle of morality through *deduction*[25]. This requires to demonstrate its unconditional bindingness for all finite rational beings (Allison 2011, 273-74). In other words, the question to be answered in section three of the *Groundwork* is: can the supreme principle of morality be justified? (Paton 1947, 199)

[25] Although Kant does not call his deduction, in section three of the *Groundwork*, a transcendental deduction since he reserves that term for theoretical concepts and principles, it has similar function, i.e., as exorcising a specter. The specter is that morality is nothing more than a chimerical idea or phantom of the brain; for this, see (Allison 2011, 274)

Kant tries to justify the moral law i.e., the principle by which rational agents as such would necessarily act if reason had full control over passion, and it is only through this that he attempts to justify it as a categorical imperative. That means, "If it is a principle on which a fully rational agent would necessarily act, it must also be on, his view, a principle on which an imperfectly rational agent ought to act, if he is tempted to do otherwise". Thus, it is possible to pass without difficulty from an unconditioned principle of action (i.e., moral law) to a categorical imperative (Paton 1947, 199).

Since the formula of autonomy describes the essence of the categorical imperative, our question can be put in the form of "how is a categorical imperative possible?" Kant postponed this question to this section because of its difficulty. This difficulty arises because a categorical imperative is a synthetic *a priori* practical proposition. Like other synthetic a priori propositions, it is necessary to find a third term that connects the subject term and the predicate of the synthetic proposition. In this section, Kant tells us what this third term is (Paton 1947, 199-200).

Thus, section three of the *Groundwork* is established on a different ground compared to the first two sections. The last two sections might be influenced by Kant's doctrine, but they at least claim to provide an analysis of the implications of our ordinary moral views. In section three, we have to enter into Kant's own metaphysical doctrines, i.e., his idea of freedom and necessity, and his distinction between the phenomenal (the sensible world) and the noumenal world (the intelligible world) (Paton 1947, 200).

3.2. The Concept of Freedom and Autonomy of the Will

We have a clear concept of autonomy, a defining property of moral will, in the last section of the *Groundwork*. Section three of the *Groundwork* is concerned with justification that human beings have a will that has this property. Kant claims that the concept of freedom provides the

key for this endeavor (G 4:446). This is because, as we have seen in the explanation of autonomy in section two of the *Groundwork*, freedom is a necessary condition for its possibility. Freedom is not merely a necessary but also a sufficient condition for autonomy and thus of morality (Timmermann 2007, 121).

Let me go to the detail of Kant's argument. Kant begins section three of the *Groundwork* with a new definition of a will. Initially, he defined will as "the power of a rational being to act in accordance with its conception of laws, i.e., in accordance with principles (G 4:412)". But, he now considers it as "a kind of causality of living beings in so far as they are rational (G 4:446)". Will is considered as a kind of causality, because it is the power of a rational being to produce effects in the world experience. However, "If we conceive the will to be free, we must mean in the first place that the will is a power to produce effects without being determined or caused to do so by anything other than itself (Paton 1947, 209)." Freedom is a quality attributed to a special kind of causality, or it is a quality that is attributed to a special kind of causal action. As such, it is opposed to natural necessity, a quality of causality attributed to non-rational beings and is described as the property of being determined or caused to act causally by alien causes (G 4:446). Non-rational beings can act causally only so long as they are caused to do so by something else other than themselves. For a billiard ball, for instance, to cause another billiard ball to move, it has to itself be caused by something else (Paton 1947, 209).

"If the will of a rational agent is conceived as free, this must mean that we regard his causal actions, or more precisely his volitions, as not determined by causes external or alien to himself (Paton 1947, 210)." The expression "alien cause" refers to, for Kant, not just external causes in the usual sense (i.e., events or entities external to the agent) that exert physical force on the agent, but also to inner causes found in the antecedent state of the agent (Allison 2011, 285).

Kant holds that since his initial characterization freedom in terms of a complete independence of natural determination is merely negative, it is not informative into its nature. However, if this negative concept of freedom is rejected, it would be impossible to justify the positive concept of freedom. In fact, the positive concept of freedom follows from the negative one (G 4:446).

To show freedom is equivalent to autonomy, Kant needs to move from the negative concept of freedom to the positive one. Having defined will as a kind of free causality, he attempts to do this by the concept of causality. Kant asserts that the concept of causality implies the concept of laws in accordance with which through something which we call 'cause' something else namely, the effect must be posited. Thus, a causality characterized not by natural necessity but by freedom cannot be lawless, but must accord with unchanging laws of a special kind. Otherwise a free will would be a logical absurdity (G 4:446). However, as H. J. Paton shows, the ground for this assertion is inadequate. Because, the law to which Kant talks about appears to be a law that connects causes and effects. Thus, we can apply this only to natural necessity. It is difficult to pass from this to a law of freedom. Because the law of freedom is a law of causal action considered in itself. (Paton 1947, 211).

According to Paton, there is a more strong force in the assertion that a lawless free will is an absurdity. However, this view is not derived from any necessary connection between causality and law. It rather comes into being from the fact that a lawless free will would be governed by chance so that it cannot be regarded as free (Paton 1947, 211).

As H. J. Paton says, if Kant's doctrine is based on the argument that presupposes the necessary connection between laws and causality, it is fallacious. Paton holds that the argument is "as superfluous as it is weak". It is because there is no need to make a connection between free will

and law on the basis of the concept of causality (Paton 1947, 211). The very definition of will, as given before, is "the power of a rational being to act in accordance with its conception of laws, i.e. in accordance with principles (G 4:412)". If so, a lawless free will is a self-contradiction; it cannot be a will at all (Paton 1947, 212).

Kant, at this stage, is not attempting to justify that a will is free. But, rather, he is just showing how a free will, if there is such a thing, would be conceived (Paton 1947, 212). It is obvious that a will governed by laws of natural necessity is not free. Because, in nature, the law governing causal action is not self-imposed but imposed by something else. This is what is called heteronomy. Thus, if a free will is not lawless, its laws must be special kinds of laws which differ from laws of natural necessity. The only way to distinguish laws of freedom from laws of nature is to suppose that laws of freedom are self-imposed. Thus, the causal action of free will is conducted by self-imposed law. But, this is what does mean by autonomy. A free will must be autonomous will. Moral law is the law of autonomous will. Thus, a free will must be conceived as acting under the moral law. Kant concludes that a free will and a will under moral law are one and the same thing. Thus, Kant holds that if we could presuppose the concept of freedom, the concept of autonomy and thus morality follows, through mere analysis, from the concept of freedom (G 4:446-47).

3.3. Freedom as a Necessary Presupposition

According to Kant, if freedom could be established in the way described before, the concept of autonomy and thereby the supreme principle of morality would follow by a mere analysis of the concept of freedom. However, to justify the moral principle, only defining freedom or describing the characteristics of freedom which must be present in a will if it is to be considered as

free is not enough. Rather, we have to show that every rational being with a will is, and indeed must be, free in the way explained before. This is because "in as much as morality serves as a law for us only in so far as we are rational beings, it must also be valid for all rational beings. And since morality must be derived solely from the property of freedom, one must also show that freedom is the property of the will of all rational beings" (G 4:447). That means, the legitimization or justification of morality depends on the connection between freedom and the will of every rational being (Allison 2011, 303).

It is useless to appeal to the experience of human action to show that the will of every rational being is necessarily free. This is because experience of freedom, if it were possible (indeed impossible), gives only a fact rather than a necessary connection between the will of every rational beings and freedom (G 4:448).

Kant argues that we can show that a rational agent as such can act on the presupposition that he is free (Under the Idea of freedom). That means, so long as we are rational beings, we necessarily act under the idea of freedom (Paton 1947, 217). The establishment or justification of this necessary presupposition would be enough to justify the moral law. If a rational agent must act on the presupposition that he is free, he must act on the presupposition that he is under moral law (Paton 1947, 218).

In order to justify the presupposition of freedom, Kant begins with theoretical reason rather than just practical reason. He claims that "We cannot possibly conceive of a reason as being consciously directed from outside in regard to its judgements." If a rational being were conscious of any such external influence, he would consider his judgement as determined, not by reason, but rather by impulse. Reason must, if it is to be reason at all, consider itself as the author of its own principles and capable of functioning in accordance with these principles independently of external

influences (G 4:448). This means, according to Kant, that reason must consider itself as free both negatively and positively in the act of reasoning (Paton 1947, 218).

Kant holds that the above argument which is valid for theoretical reason must equally be valid for practical reason (i.e., as a rational will or as reason exercising causality) (G 4:448). That means, a rational agent as such must, in action, presuppose his/her rational will to be the source of its own principles of action and to be capable of functioning in accordance with these principles. To say the same thing in other words, a rational agent must, in action, presuppose his/her rational will to be free both negatively and positively; to be free from any alien or external determination and free to obey its own rationale principles. To say this is tantamount to that a rational being must act only on the presupposition that he/ she is free (i.e., he/she must act under the Idea of Freedom). Only then can a rational being considers his/her will as his/her own. This is the doctrine that Kant wants to establish and from which the principle of morality follows analytically (Paton 1947, 219).

H. J. Paton argues that, for Kant, freedom is a necessary presupposition for both all actions and thinking. This means that a rational being can act, just as he/she can think, merely on the presupposition of freedom. A rational being implicitly presupposes freedom both in his/her action and thinking. If not, there is no anything as action and there is no such thing as will (Paton 1947, 219).

3.4. The Vicious Circle

Kant's argument seems complete. But, he worries that his argument for human autonomy might contain a vicious circle. Morality was 'traced back' to the idea of freedom at the first sub-section of section three of the *Groundwork*, and in the next subsection it was argued that freedom needs to be necessarily presupposed if we are to think of ourselves, as rational agents, bound by

moral laws. Thus, we have not yet been presented with an independent ground of the idea of freedom (G 4:448).

3.5. The Two Standpoints

Kant introduces us to the doctrine of the two standpoints as a way out of the vicious circle. This doctrine of the two standpoints is an essential element of his Critical Philosophy. In dealing with it, he neither assumes that the whole argument of the *Critique of Pure Reason* to be familiar to his readers of the *Groundwork* nor he attempts to repeat the detail of these arguments in a short treatise on ethics. He just gives us the highlight of his doctrine of metaphysics which is insufficient, strictly speaking, taken by themselves (Paton 1948, 43).

According to Kant, all the ideas which are provided to our senses come to us without our own volition or come to us involuntarily; that is, we are passive. We assume that these ideas come to us from objects. By means of these ideas, we can know objects only as they affect us. However, we do not know what these objects are in themselves. This leads to the division between things as they appear to us (appearance) and things as they are in themselves (things in themselves). We know only appearances. However, we have to assume that behind appearances, there are things in themselves even if we do not know these things in themselves, but as they affect us. This gives rise to a rough distinction between a sensible world and an intelligible world. The world of sense is given through sense and varies in accordance with the difference of sensibility in various observers. On the other hand, the intelligible world can be conceived but never known since knowledge requires sensing and conceiving and it remains always the same (G 4:451).

Kant also argues that this distinction applies to man's knowledge of himself. Man can know himself only as he appears by means of inner sense or through introspection. However, behind

this appearance, he must assume, there is an Ego as it is in itself. In so far as he is known through introspection or inner sense, and in so far as he is capable of receiving sensations passively, he must consider himself as members of the sensible world. On the other hand, in so far as he is capable of pure activities without the influence of sense, he must consider himself as members of the intelligible world. The intelligible world is called the intellectual world; the world which is intelligible since it is intelligent. But, Kant suggests that we know nothing about this world (G 4:451).

Kant claims that man really finds in himself a pure activity which is free from the influence of sense. That is, he finds in himself a faculty of reason. The power of reason distinguishes man from any other things even from himself as long as he is affected by objects (G 4:452). Here, as Paton noted, Kant appeals to theoretical reason, as he did before. But, he now uses it in his own critical sense (Paton 1948, 44). According to Kant, we have a spontaneous faculty of understanding. The power of understanding along with other factors produces from itself concepts or categories and it uses these concepts or categories to bring the ideas of sense under rules. Therefore, although the faculty of understanding is genuinely spontaneous, it is still bound up with sense. Without the use of sensibility, it does not do anything at all. On the other hand, reason is a power of ideas. That means, it produces unconditioned concepts that goes far beyond what sensibility can offers. Reason, unlike understanding, can show a pure spontaneity which is totally independent of sense (G 4:452).

In virtue of this spontaneity, Kant argues that man must conceive himself, *qua intelligence*, as members of the intelligible world and regards himself as subject to laws that are grounded merely in reason. On the other hand, as long as man is sensuous and known to himself through inner sense or introspection, he must regard himself as belong to the world of sense (the sensible

world) and regard himself as subject to laws that have their grounds in nature alone. According to Kant, these are the two equally legitimate standpoints from which man (a finite rational being) must regard himself (G 4:452).

For Kant, the above doctrine of theoretical reason is equally applicable to pure practical reason. That means, since man, as a finite rational being, from one standpoint, can regard himself as belonging to the intelligible world, he can regard his will as free from the determination of sensuous causes and as obedient to laws grounded merely in reason (G 4:452). This is tantamount to saying "As a rational being, and thus as a being belonging to the intelligible world, human being can never think of the causality of his own will otherwise than under the idea of freedom; for, independence from the determining causes of the world of sense (which reason must always ascribe to itself) is freedom (G 4:452)". From this, there follows the principle of morality (the moral law) and the categorical imperative (G 4:452). As we understand from the above argument, Kant's argument proceeds from being member of the intelligible world to the idea of freedom.

The suspicion of the vicious circle, according to Kant, is now avoided. Because "we now see that when we think of ourselves as free we transfer ourselves into the intelligible world as members and cognize autonomy of the will along with its consequence, morality (G 4:453)". But, when man regards himself as members of both the intelligible and the sensible world, he can recognize the moral law as a categorical imperative (G 4:453). Kant, here, argues from freedom to the intelligible world, which is directly opposed to the previous argument.

As we understand from the argument presented so far, in the *Groundwork*, Kant does not clearly show whether he infers from the concept of freedom to the intelligible world or from being member of the intelligible world to the concept of freedom, or whether he is establishing a

reciprocal connection between being members of the intelligible world and the concept of freedom. For sure, if Kant uses the last option, he does not reject the vicious circle.

3.6. How Is a Categorical Imperative Possible?

In order to answer this question, Kant begins again by maintaining that a rational being must regard himself as belonging to the intelligible world. From this, Kant infers that a rational being considers himself/herself as exercising causality and manifesting a free will. According to Kant, man, as a finite rational being, must also regard himself from the standpoint of the sensible world. In connection to this, Kant argues that if I were solely a member of the intelligible world, all my actions would necessarily conform to the principle of autonomy; if I were solely a member of the world of sense, they would necessarily be exclusively subject to the law of nature (G 4:453). At this point, as H. J. Paton properly mentions it, Kant inserts a strange argument and at the same time confused in expression and difficult to interpret (Paton 1948, 45): "The intelligible world contains the ground of the sensible world and therefore also the ground of its laws; consequently, the intelligible world is (and must be thought of as) directly legislative for my will (which belongs wholly to the intelligible world) (G 4:453)." From this premise which is itself problematic in the sense that it needs a considerable expansion, Kant infers that the law governing my will as a member of the intelligible world (from one standpoint) ought to govern my will although I am also, from another standpoint, a member of the world of sense (G 4:453); "I must regard the laws of the intelligible world as imperative for me and the laws conforming to this principle as duties (G 4:454)." As we understand from his word, Kant here introduces a metaphysical argument from the superior reality of the intelligible world and the rational will.

Kant concludes his argument by saying that "categorical imperatives are possible because the idea of freedom makes me a member of an intelligible world. Now if I were a member of only

that world, all my actions would always accord with autonomy of the will (G 4:454)." But, because I am also a member of the world of sense, I experience the moral law as an imperative or my actions ought to accord to it (G 4:454). This categorical "ought", Kant tells us, presents a synthetic *a priori* proposition. Thus, to connect this ought with my finite will and by extension the will of any other rational being like me, a third term is needed. In other words, to establish the thesis that the will of every finite rational being ought always to act in accordance with the moral law, it is needed to combine the following two separate concepts: (i) the human will which is affected by sensuous desires and inclinations and (ii) the principle morality. The third term that connects these two different concepts is the same will but viewed as *a pure will* belonging to the intelligible world. Consequently, the laws of the intelligible world (i.e., autonomy) must be the condition of actions of human will which is affected by sensuous desires (G 4:454).

However, the deduction that Kant, in section three of the *Groundwork*, assumes to be successful is problematic. To show the deduction or the justification of a categorical imperative is unsuccessful, it is crucial to consider at least three problems mentioned by Jens Timmermann. (Timmermann 2007, 142). I shall return to this issue later.

Chapter Four

Conclusion in Critique Form

In the *Groundwork*, the aim of Kant is to search for and establish the supreme principle of morality. Categorical imperative is the supreme principle of morality in its imperative mood. However, Kant does not insist that he discovered the categorical imperative. This is because the supreme principle of morality has implicitly been used by any rational agents in order to make their choice and judgment. Thus, he aimed at exposing the moral principle which has been implicitly used by the pre- philosophical understanding of morality. Although in the pre-philosophical understanding of morality any ordinary good men used this principle to judge moral actions, they were unable to make this principle clear for themselves.

Accordingly, in section one of the *Groundwork*, Kant begins with ordinary ways of thinking about morality and moves into exposing the principle that ordinary good man is supposed to implicitly use in judging his action. Kant's intention is to derive it through the analysis of the concept of a good will. This is because any person who does the right thing for the right reason evinces what we call a good will. Thus, through analysis of the concept of a good will, which is considered by common human reason to be the only good without qualification, Kant reaches a version of the categorical imperative as the principle in which a good will under human condition would act.

Kant holds that his analysis fully agrees with ordinary human reason. Although the ordinary good man does not establish this moral principle in abstraction, he really uses it in evaluating particular moral matters. Even in practical matters, ordinary human reason is more important than philosophy.

But, common moral cognition is pre-philosophical in its origin and it is a species of innocence in its unreflective form. Thus, it is incapable of protecting itself against evil. It will be easily seduced. Therefore, it is subject not only to philosophical explication, but also to rigorous philosophical criticism, correction, and rejection.

In section two of the *Groundwork*, Kant continues to expose the supreme principle of morality but now through different way. He does not start from the analysis of common moral cognition but rather he explores it from a philosophical stand point. Kant continues to expose the supreme principle of morality in section two of the *Groundwork*, in order to show that this principle has the characteristics of absolute necessity. This cannot be shown by claiming that it is implicitly found in common human reason. Because this may imply that this principle is grounded in human nature which in turn prevents it from possessing the characteristics of absolute necessity.

In section two of *Groundwork*, Kant begins his new argument with some analysis of rational beings. All things in nature act in accordance with laws, but only rational beings act according to the conception of a law. That is, it is only rational beings that have the capacity to recognize and consult laws in order to guide their actions. The laws according to its representation finite rational beings act bring about the concept of imperatives that constrain the will.

Kant suggests all imperatives command either hypothetically or categorically. To command hypothetically is to command something under certain condition. It is a command to achieve a certain presupposed end. By contrast, to command categorically is to command unconditionally, or independent of any presupposed end.

For Kant, since the moral law, if it exists, must apply universally and necessarily, it cannot be based on hypothetical imperatives. The imperative related to the moral law must be a categorical. Since the categorical imperative applies to all rational beings regardless of the various ends a

person would have, it could be the basis of the moral law. As we know, a categorical imperative could not be based on a particular end. Thus, Kant holds that the categorical imperative must be based on the notion of a law itself. Laws by definition, apply universally. From this, Kant derives the content of the categorical imperative that requires moral agents act only in a way that the maxim of their action can be a universal law. And we call it formula of universal law (FUL). The categorical imperative, as I show before, is Kant's general expression of the supreme principle of morality in its imperative mood, but Kant continues to provide three different formulas of this general formulation. These additional formulas refer to different stages that in turn enable him to fully construct the concept of the categorical imperative.

The first formula of the categorical imperative is the formula of law of nature (FLN). This formulation states that an action is only morally permissible if every moral agent can adopt the same maxim of action in a world in which he/she is a member without contradiction.

The second formulation of the categorical imperative is the Formula of Humanity (FH). Kant reaches this formula by considering the motivating ground of the categorical imperative. Since the moral law is necessary and universal, its motivating ground must have absolute worth. If we find something with an end in itself, it would be the only possible ground of a categorical imperative. Kant asserts that every rational being exists as an end in itself.

The third formulation of the categorical imperative is the formula of autonomy.This formula is formulated on the basis of the idea that a rational will makes itself the law that it obeys. It takes important elements from both the mentioned two formulas. The first Formula (FLN) specifies the universality of laws, while the second formula (FH) is more subjective and focuses on how

you treat the person with whom you are interacting. Thus, by combining the objective and subjective aspect of the two formulas, Kant leads to the idea that every rational being is involved in making universal laws.

Kant suggests that the notion of autonomy further leads to another closely connected world, namely the realm of ends. When one considers himself as an autonomous agent, he at the same time thinks of himself as a law-giving member of the realm of ends. Kant defines realm as a systematic union of different rational beings under common laws. The term systematic refers to the various members that live harmoniously and support one another.

In relation to the realm of ends, Kant reformulates the formula of autonomy in the shape of formula of realm of ends (FRE): *act in such a way that your maxim could be a law in the realm of ends* (cf. G 4:439). Since FRE can be derived by combining the ideas of previously mentioned formulas of the categorical imperative, it could be more adequate in expressing the spirit of the supreme principles of morality.

Kant asserts that FLN, FH and FA/FRE are not independent ethical theories. They are rather reformulations or variants of the single categorical imperative. The single categorical imperative represented by these formulas is the FUL. Thus, in saying this, Kant accomplishes the first of the second task that he aims to fulfill in the *Groundwork,* namely formulating the supreme principle of morality.

In section three of the *Groundwork,* Kant attempts to accomplish the second project that he aims to fulfill in the *Groundwork,* namely to establish the objective validity of the supreme principle of morality through transcendental *deduction.* In the last two sections of the *Groundwork,* Kant claims that he is not maintaining the truth of morality. What he has done is that he has determined the condition of accepted moral beliefs through analytic argument. If

so, it is impossible to justify ordinary moral beliefs progressively from this ultimate precondition without a vicious circle. Thus, Kant has to establish these ultimate preconditions independently through transcendental deduction. That is, it is required a transcendental deduction of the supreme principle of morality. Its function is to establish the possibility that moral judgements are valid. At present, Kant insists that it is an open question whether morality and its condition on which it is founded are not merely phantom of the brain or illusions.

To show that this worry is groundless, in section three of the *Groundwork,* Kant begins with a new definition of a will. He considers will as a causality of rational beings. Will is considered as a kind of causality, because it is the power of a rational being to produce effects in the world experience. But, if the will were determined by laws of nature, it would not be free. Thus, a free will must be conceived as one free from any external influence.

This is only negative concept of freedom. But, this negative concept of freedom is important to know the positive one. The positive concept of freedom is autonomy. He leads to the positive concept of freedom through the concept of causality. Kant asserts that a causality characterized by natural necessity is governed by laws. Thus, a causality characterized not by natural necessity but by freedom cannot be lawless because a lawless free will would be a logical absurdity. Thus, a free will must act under laws. However, the way he asserts this is inadequate. Because the law to which Kant talks about is a law that connects causes and effects. Since we can apply this only to natural necessity, it is difficult to pass from this to a law of freedom. Because the law of freedom is a law of causal action considered in itself. From this problematic premise, he argues that the laws of free will could not be one imposed upon it by others. If it is not other imposed, it must be self-imposed. But, this is what does mean by autonomy. Thus, *a free will is equivalent to autonomous will*. Since moral law is the law of autonomous will, a free will must be conceived as acting under

the moral law. Kant concludes that a free will and a will under moral law are one and the same thing. Thus, Kant holds that if we could presuppose the concept of freedom, the concept of autonomy and thus morality follows, through mere analysis, from the concept of freedom.

However, if Kant's doctrine is based on the argument that presupposes the necessary connection between law and causality, it has to be rejected as fallacious. As a result, Kant fails to connect autonomous will and free will which is one of two connections he wants to fulfil in the *Groundwork.*

According to Kant, since morality is valid for all rational beings, justification of morality depends on the connection between freedom and the will of every rational being. Although it is impossible to show this through experience, it is possible to show that a rational agent as such can act on the presupposition that he is free. That means, so long as we are rational beings, we necessarily act under the idea of freedom. The justification of this necessary presupposition would be enough to justify the moral law.

Kant's argument seems complete. But, as I have shown in the last chapter, he worries that his argument presented so far might contain a vicious circle.

As a way out this vicious circle, Kant introduces the doctrine of the two standpoints. In the *Critique of Pure Reason*, he suggests that all ideas which are provided to our senses come to us from objects. Through these ideas, we know objects only as they affect us, but not in themselves. This leads to the division between appearance and things in themselves. This in turn leads to the distinction between a sensible world and an intelligible world. The world of sense is given through sense. But, the intelligible world can be conceived but never known.

Kant applies this distinction to man's knowledge of himself. Man can know himself only as he appears through introspection. But, behind this appearance, he must assume, there is an

Ego as it is in itself. In so far as he is known through introspection, he must consider himself as members of the sensible world. On the other hand, so far as he is capable of pure activities without the influence of sense, he must consider himself as members of the intelligible world.

Kant claims that man really finds in himself a faculty of reason which is pure activity and free from the influence of sense. Kant here is talking about theoretical reason and applies this to practical reason. Thus, through the faculty of reason, man must conceive himself as members of the intelligible world and regards himself as subject to laws that are grounded merely in reason. On the other hand, so far as man is sensuous and known to himself through inner sense, he must regard himself as members of the sensible world and as subject to laws grounded merely in nature. These are the two equally legitimate standpoints from which man must regard himself.

The doctrine of theoretical reason is equally applicable to pure practical reason. Thus, since a finite rational being, from one standpoint, can regard himself as members of the intelligible world, he can regard his will as free from any sensuous determination and as obedient to laws grounded merely in reason. From this, there follows the principle of morality and the categorical imperative. Here, Kant argues from membership of the intelligible world to the idea of freedom although he rejects it in the next argument.

To show that the suspicion of the vicious circle is avoided, Kant argues from freedom to the intelligible world, which is directly opposed to the previous argument and that shows his inconsistency and hesitation in the *Groundwork*. He says, more specifically, that *"*when we think of ourselves as free we transfer ourselves into the intelligible world as members and cognize autonomy of the will along with its consequence, morality (G 4:453)". But, when man regards himself as members of both the intelligible and the sensible world, he can recognize the moral law as a categorical imperative.

Thus, as it has been shown above, in the *Groundwork*, Kant does not clearly show whether he infers from the concept of freedom to the membership of the intelligible world or from being member of the intelligible world to the concept of freedom, or whether he is establishing a reciprocal connection between being members of the intelligible world and the concept of freedom. For sure, if Kant uses the last option, he does not reject the vicious circle. Kant also repeats this inconsistency or hesitation when he attempts to answer the question "how is a categorical imperative possible?" as it is shown below.

To answer the question "how is a categorical imperative possible?", he begins by maintaining that "a rational being must regard himself as belonging to the intelligible world, only then a rational being considers himself as exercising causality and manifesting a free will (G 4:453)". A finite rational being must also regard himself from the standpoint of the sensible world. At this point, Kant inserts a strange argument and at the same time confused in expression and difficult to interpret: "The intelligible world contains the ground of the sensible world and therefore also the ground of its laws; consequently, the intelligible world is (and must be thought of as) directly legislative for my will (which belongs wholly to the intelligible world) (G 4:453)." From this problematic premise, Kant infers that the law governing my will as a member of the intelligible world ought to govern my will although I am also, from another standpoint, a member of the world of sense. In this problematic argument, Kant seems to introduce a metaphysical argument from the superior reality of the intelligible world and the rational will. If he really uses a metaphysical doctrines to establish the principle of morality or categorical imperative, indeed he does, Kant commits a fundamental error since it is impossible to deduce moral obligation from metaphysical considerations which has nothing to do with morality. In fact, Kant himself seems to be aware of his mistake. The reason is that, at the preface of the *Groundwork*, Kant holds that he has a plan to publish

a book entitled *Metaphysics of Morals*. The *Groundwork* serves merely as a preliminary to *Metaphysics of Morals*. But, he first publishes *Critique of Practical Reason* rather than *Metaphysics of Morals*, and some of arguments in the *Critique of Practical Reason* are different from the *Groundwork*. This implies that Kant himself seems to be aware of that in the *Groundwork* he has no sufficient reason to establish the categorical imperative.

As we have seen above, to answer the question "how is a categorical imperative possible?" he argues from membership of the intelligible world to freedom. But, when he concludes his argument in asserting the possibility of the categorical imperative, he argues from freedom to membership of the intelligible world in such a way that "categorical imperatives are possible because the idea of freedom makes me a member of an intelligible world. Now if I were a member of only that world, all my actions would always accord with autonomy of the will (G 4:454)." But, because I am also a member of the world of sense, I experience the moral law as an imperative or my actions *ought* to accord to it. This categorical "ought" presents a synthetic *a priori* proposition. Thus, to connect this ought with every finite will, or to establish the thesis that the will of every finite rational being ought always to act in accordance with the principle of morality, it is needed to combine the following two separate concepts: the human will or finite will and the principle morality. The third term that connects these two different concepts is the same will but viewed as *a pure will* belonging to the intelligible world. Consequently, the laws of the intelligible world must be the condition of actions of human will which is affected by sensuous desires.

In addition to problems mentioned before, there is another fundamental error that shows the deduction that Kant, in section three of the *Groundwork*, assumes to be successful is failed. That means, as I mentioned before, Kant, in the *Groundwork*, attempts to justify the bindingness

of the categorical imperative for finite will through deduction more specifically through transcendental deduction. He ends this project in the way presented so far. But, it is unsuccessful. To show this, it is important to state at least three problems mentioned by Jens Timmermann (2007, 142). First, Kant claims that the two concepts which are supposed to be connected in a synthetic judgement must be connected by a third element that contains both of them. But, in the above case, the idea of a pure will located in the intelligible world may contain the laws of autonomy (the moral law), but it is not obvious that it contains the idea of itself as a finite will or a human will. Secondly, Kant, in the *Critique of Pure Reason*, thinks that the minimum condition for the objective validity of synthetic judgment is *time*. That means, a synthetic judgement cannot be objectively valid if the 'third something' is not provided in intuition of time. In light of this, the deduction or the justification of the categorical imperative in the *Groundwork* is unsuccessful, since a pure will is not provided in intuition of time. Thirdly, some interest in moral behavior is needed to conduct the connection in practice: reverence for the moral law. But, the possibility of explaining this interest lies beyond the boundaries of practical philosophy as Kant himself admits.

In general, arguments that Kant uses, in the *Groundwork*, to justify the binding character of the categorical imperative for finite rational beings are problematic, obscure and inconsistent. As we have seen so far, some of the arguments in the *Groundwork* contradicts with each other and others contradict with arguments in the *Critique of Pure Reason*. Thus, the existence of these problems implies that Kant, in the *Groundwork*, does not have a sufficient reason to establish the categorical imperative. In other words, Kant, in the *Groundwork*, does not put the categorical imperative or morality on a solid ground.

References

Allison, Hennery E. 2011. *Kant's Groundwork for the Metaphysics of Morals: A commentary.* Cambridge: Cambridge University press.

Ellington, James W. 1993. Introduction to *Grounding for the Metaphysics of Morals.* Translated by James W. Ellington. Indianapolis: Hackett publishing company.

Kant, Immanuel. 1998. *Groundwork of the Metaphysics of Morals*. Translated and edited by Mary Gregor. Oxford: Oxford University Press.

____1993. *Grounding for the Metaphysics of Morals: With on a Supposed Right to Lie because of Philanthropic Concerns*. Translated by James W. Ellington. Indianapolis: Hackett pub lishing company.

_____1991. *The Metaphysics of Morals*. Translated by Mary Gregor. New York: Cambridge University press.

_____1998. *Critique of Pure Reason*. Translated and Edited by Paul Guyer and Allen Wood. Cambridge: Cambridge University Press.

_____2002. *The Critique of Practical Reason*. Translated by Werner S. Pluhar. Indianapolis/Cambridge: Hackett publishing company, Inc.

___1997. *Lecture on Ethics*. Translated by Peter Heath. Cambridge: Cambridge University Press.

___1948. *The Moral Law or Kant's Groundwork of the Metaphysic of Morals*: *A New Trans lation with Analysis and Notes*. Translated by H.J. Paton. London: Hutchinson's University library.

____200. *First Introduction to the Critique of the Power of Judgment*. Translated by Paul Guyer. Cambridge: Cambridge University Press.

Korsgaard, Christine M. 1998. Introduction to *Groundwork of the Metaphysics of Morals*. Translated and edited by Mary Gregor. Oxford: Oxford University Press.

___1996. *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press

Paton, H. J. 1947. *The Categorical Imperative: A Study in Kant's Moral Philosophy*. Philadelphia, Pennsylvania: University of Pennsylvania Press.

____1948. Introduction to *the Moral Law or Kant's Groundwork of the Metaphysic of Morals*: *New Translation with Analysis and Notes*. Translated by H. J. Paton. London: Hutchinson's University library.

Timmermann, Jens. 2007. *Kant's Groundwork of the Metaphysics of Morals: A commentary*. Cambridge: Cambridge University press.

Willaschek, Marcus. 2006. "Practical Reason", in *Groundwork for the Metaphysics of Morals* edited by Horn and Schönecke. Berlin: Walter de Gruyter.

Wood, Allen W. 1999. *Kant's Ethical Thought*. Cambridge: Cambridge University Press.

___.2008. *Kantian Ethics*. Cambridge: Cambridge University Press.