

**KANT'S MORAL IDEALISM: THE LOGICAL BASIS AND METAPHYSICAL
ORIGIN OF THE IDEAS OF COMMUNITY AND AUTONOMY**

Lucas Thorpe

A DISSERTATION

in

Philosophy

Presented to the Faculties of the University
of Pennsylvania in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy

2003

Dissertation Supervisor

Graduate Group Chairperson

Kant's moral idealism – an examination of the metaphysical and logical basis of the ideas of community and autonomy.

Copyright

Acknowledgements (with page number)

ABSTRACT

KANT'S MORAL IDEALISM: THE LOGICAL BASIS AND METAPHYSICAL ORIGIN OF THE IDEAS OF COMMUNITY AND AUTONOMY

Lucas Thorpe

Paul Guyer

This thesis examines the theoretical foundations of Kant's moral philosophy. I argue that Kant's moral ideal of a kingdom of ends is to be identified with the theoretical idea of a community, and that this idea can be traced back to the category of community introduced in his table of categories. In particular I argue that, for the mature Kant, (a) the only application of the theoretical idea of community is the moral idea of a kingdom of ends, (b) the only way we can conceive of a kingdom of ends is as a political community governed by juridical laws, and (c) the only way we can conceive of a member of a community is as an autonomous agent.

Table of Contents

Introduction	1
Part One - Ethics	
Chapter 1 - The Idea of a Good Will and the Value of Moral Philosophy	18
Introduction	18
(1a) Moral Constructivism & Moral Idealism	27
(1b) Kant's Strategy in the <i>Groundwork</i>	33
(1c) Virtue and Holiness	39
(1d) Kant's Disagreement with Baumgarten & Wolff	47
(i) Rationalism	48
(ii) Perfectionism	55
(1e) An Objection to My Interpretation	58
(1f) The Value of Moral Philosophy	63
Chapter 2 - Moral Idealism and Kant's Ethics of Interaction	83
(2a) Kant's Moral Idealism	83
(2b) Kant's Phenomenal Solipsism	100
Chapter 3 - The Kingdom of Ends as a Community of Spirits: Kant and Swedenborg	112
Introduction	116
(3a) Kant's Change of Position	117
(3b) 1763-6 - Kant's Engagement with Swedenborg	121
(3c) Swedenborg	123
(3d) Kant's Encounter with Swedenborg: The Facts	131
(3e) Swedenborg's Heaven and Kant's Ideal of a Moral Community	142
Chapter 4 - From Pneumatic Laws to Moral Laws: Interaction and Autonomy	165
(4a) Pneumatic Laws and Forces in <i>Dreams</i>	167
(4b) From Pneumatic Laws to Moral Laws	172
(4c) Moral Laws & Autonomy	175
(4d) Real Interaction is Only Possible Between Autonomous Agents	179
(4e) Conclusion	184

Part Two: Metaphysics

Chapter 5 - The Critical Kant and Traditional Metaphysical Enquiry	186
(5a) Introduction	187
(5b) Kant's Critique of Traditional Metaphysics	190
(5c) The Positive Role of Traditional Metaphysical Speculation	197
(5d) The Conditioned Nature of all Objects of Experience	213
Chapter 6 - Theories of Interaction: An Historical Overview	224
(6a) Introduction	224
(6b) The Three Standard Accounts of Interaction	227
(6c) Kant's Rejection of Pre-established Harmony	230
(6d) Kant's Attitude Towards Physical Influx	234
Chapter 7 - Kant's Theory of Action	241
(7a) Introduction	241
(7b) A Problem	244
(7c) Kant's Solution	246
(7d) Leibniz' Problem in Explaining the Possibility of Composite Substances	251
Chapter 8 - The Idea of Community and the Disjunctive Form of Judgment	258
(8a) The category of community	260
(8b) The Distinction Between the <i>Category</i> of Community and the <i>Idea</i> of a Community	267
(8c) The Distinction Between the Ideas of a Real and an Ideal Community	276
 Conclusion 	
Chapter 9 - The Civil Condition as a Real Community: An Analysis of Kant's Ontology of Property	279
(9a) Introduction	281
(9b) The Ontology of Property	287
(9c) The Doctrine of Right and the Doctrine of Virtue	298
(9d) The Kingdom of Ends is the Idea of a Political State	306

Introduction

Many scholars working on Kant's ethics have been influenced by Rawls's claim that Kant should be read as a **moral constructivist**. According to this interpretation, our moral ideas, such as the idea of a kingdom of ends, are constructed by practical reason, which is understood to be a faculty entirely distinct from theoretical reason. Practical reasoning is governed by what Rawls and his followers call the Categorical Imperative (CI) Procedure, and our moral ideas are somehow constructed using this procedure. On such an interpretation the categorical imperative (procedure) is logically prior to the idea of a kingdom of ends, and we can investigate the nature of this procedure without engaging with Kant's theoretical philosophy. Rawls' interpretation has been extremely influential, and many of his former students are now prominent Kant scholars. One result of this influence has been the practice of studying Kant's ethics divorced from his theoretical work. I believe that such a divorce has led to an impoverished understanding of Kant's ethics.

I reject the claim that Kant is a moral constructivist, instead I maintain that he should be understood as a **moral idealist**. I argue, in contrast to Rawls and his followers, that our moral ideas are theoretical ideas, and that they are independent of, and logically prior to,

the categorical imperative.¹ In particular I argue that the idea of a kingdom of ends is a particular theoretical idea, namely the pure idea of a community, which can be traced back to the category of community introduced in the table of categories of the *Critique of Pure Reason*.

As a result of this, the idea of a *member* of a kingdom of ends is also a theoretical idea; it is the idea of a member of a community. This theoretical *idea*, however, presents itself to us practically as an *ideal*, that is as something of extreme value that we should strive to instantiate. The categorical imperative is merely the practical recognition by our conscience of the theoretical idea (of a member of an ideal community) as an ideal. For the categorical imperative commands us to strive to be a member of a kingdom of ends. Kant is a moral idealist and, in contrast to the moral constructivist, believes that our moral ideas are logically prior to the categorical imperative, for the command “be a member of a kingdom of ends!” presupposed that we understand the idea of a kingdom of ends.

In the Groundwork Kant argues that to be virtuous is to strive to be a member of an ideal kingdom of ends, and I argue that the idea of a kingdom of ends is the idea of an ideal political community, governed by juridical laws. The reason for this is because the idea of a kingdom of ends is the idea of a community, and the idea of community is an idea of pure reason, being derived from the category of community, the third category of relation. For Kant the idea of community is the idea of a whole the parts (or members) of

¹ And would be logically prior to the CI Procedure if Kant believed there was such a thing.

which are simple (indivisible) and logically prior to the whole, and which mutually limit and resist one another. I explain why the idea community has these features by examining Kant's logical and metaphysical works. The fact that the idea of community is defined in these terms implies two things: (a) that the idea of a member of a community can only be the idea of an autonomous individual, and (b) our idea of a community is the idea of a political community governed by juridical laws.

(a) As a whole, our idea of a community must be unified, and if the whole is to be a *real*, as opposed to an ideal, whole the source of the unity must be intrinsic to the community. And as a community consists merely of its members, which are logically prior to the whole, the members of a real community must be the source of the unity of the community. Now, Kant believes that, a community can only be unified by laws, and so in a real community the members of the community must be the source of the laws that provide the community with its unity, that is the members of a real community must be *autonomous*.

(b) The reason our pure idea of a real community must be the idea of a political community governed by juridical laws is because a community must not only be unified, but the members of the community, if they are to be capable of real interaction, must resist one another, for Kant believes that inter-substantial action can only be understood in terms of the withdrawal of resistance, and the source of the unity must also be the source of this resistance. Now, what unifies a community are laws given by the members of the community, and so these laws must also be the source of the resistance between the

members of the community. The only type of laws, however, that can create resistance between individuals are juridical laws. Juridical laws assign rights to use objects, and these rights imply corresponding duties in others to not interfere with, or resist, my use of these objects. Such laws give individuals a legitimate right to resist others use of objects. This is why our pure idea of a real community must be the idea of a political community governed by juridical laws. Juridical laws, then, are the source of intelligible resistance, and this is the only possible basis of intelligible interaction between individuals.

This dissertation is divided into two parts. In **part one** (chapters one to four) I examine the role of the concept of community in Kant's ethical work. In **part two** (chapters five to eight) I examine Kant's theoretical account of the concepts of community and interaction. In my final chapter (chapter nine) I return to Kant's practical philosophy and show how this understanding of Kant's theoretical account of community and interaction can help us understand Kant's theory of property as laid out in the *Metaphysics of Morals*.

PART ONE - ETHICS

In **chapter one** I focus upon explaining and justifying two claims: firstly, that Kant should be understood as a moral idealist and, secondly, that, for Kant, the primary function of moral philosophy is to arrive at a clearer theoretical understanding of what it is we are striving to be insofar as we recognize the demands of morality. The chapter is divided into seven sections. In **section one** I compare my approach with that of Rawls and his followers who argue that Kant should be understood to be a *moral constructivist*.

The moral constructivists argue that our moral ideas, including the idea of a kingdom of ends, are practical (as opposed to theoretical) ideas that are constructed on the basis of a practically rational procedure. In **section two** I explain my interpretation of Kant's strategy in the *Groundwork*. In opposition to the moral constructivists, who argue that Kant's strategy in the *Groundwork* is to discover the implicit procedure of practical rationality (the "CI-procedure") and then show how the idea of a kingdom of ends can be constructed by means of this procedure, I argue that Kant's goal in the *Groundwork* is to arrive at a clearer theoretical understanding of what it is that our conscience demands that we become, and that his strategy is to begin by examining our pre-philosophical moral intuitions to show that what the person who listens to their conscience is striving to be is a member of an ideal community. Such a person does not necessarily understand theoretically that this is what they are doing, and Kant's goal in the *Groundwork* is to provide such a person with a clearer theoretical understanding of what they are attempting to be when they listen to the judgment of their conscience. In **section three** I explain Kant's account of the relationship between the notions of virtue and holiness. I argue that Kant believes that it is our duty to be holy, and that to be virtuous is, by definition, to strive to be holy. Given Kant's belief that ought implies can, the fact that we have a duty to be holy implies that it is possible for us to be holy. I explain how the postulates of practical reason introduced in the *Critique of Practical Reason* are intended to explain this possibility. In **section four** I examine Kant's disagreements with Wolff and Baumgarten, two of his rationalist predecessors who advocated a rationalist and perfectionist morality. I argue that Kant himself should be understood as being both a rationalist and a perfectionist, and that what he objects to is their *formalistic* formulation

of rationalism and perfectionism, and I suggest that he would reject moral constructivism on similar grounds. In **section five** I examine a number of passages in the *Critique of Practical Reason* that seem to offer support for a moral constructivist reading of Kant. In particular Kant, famously, insists that it is “the moral law that first determines and makes possible the concept of the Good” (5:64). I argue that these passages do not have the implications they seem to have at first sight, for after making such remarks, Kant makes it clear that this priority is merely methodological and not ontological, for he believes that although the categorical imperative is the *ratio cognoscendi* of the idea of being a member of a kingdom of ends, the idea of being a member of a kingdom of ends is the *ratio essendi* of the categorical imperative. In the **final section** of the chapter, I examine Kant’s account of the *value* of engaging in moral philosophy. Kant is not an intellectualist, for he believes that to be virtuous all we have to do is listen to our conscience, so a clearer theoretical understanding of what our conscience demands of us is not necessary in order to be moral. Given his rejection of intellectualism, however, it is not immediately clear why arriving at a clearer theoretical understanding of what morality demands has any practical value. I argue that the theoretical clarification that moral philosophy provides has motivational benefits. The reason for this is because immorality is always the result of choosing not to listen to the judgment of our conscience, and this happens, Kant believes, when we listen to excuses for not living up to the moral ideal, and these excuses are themselves the product of theoretical speculation. A clearer theoretical understanding of the moral ideal will not, in itself, make us more moral, but it will make it harder for us to give ourselves excuses for not living up to the ideal. Engaging in moral philosophy, then, quietens the excuse giving voice, the

voice of the defense attorney in the court of conscience, which makes it easier for us to be attentive to the voice of the judgement of conscience.

In **chapter two** I argue that, for Kant, to be virtuous is to strive to be a member of a kingdom of ends understood as the idea of a world or community. I argue, (a) that Kant's moral ideal is the (pure) idea of being a member of a world, (b) that our idea of a world is the idea of a *community* of individuals in interaction with one another and, (c) that the only way of conceiving of a community of individuals as in interaction is if we conceive of each individual member of the community as autonomous.

The (pure) idea of a world could also be called an intelligible world, so I am suggesting that, for Kant, the moral ideal is the idea of being a member of a/the intelligible world. In talking of an 'intelligible world', at least in the ethical context, Kant is not referring to some other, ontologically distinct world. Instead, when he talks of the idea of an intelligible world all he really means is our pure idea of a world, that is our idea of a world insofar as it is an object of the pure intellect.

I also argue in this chapter that Kant should be understood as offering an 'ethics of interaction'. I call Kant's ethics an 'ethics of interaction' because Kant's moral ideal is the idea of being an individual member of a world of individuals, and, for Kant, a world is only really a world, as opposed to a mere aggregate, if the individual beings that constitute the world are in interaction.

In **chapter three** I examine the genesis of Kant's moral theory. I argue that Kant first started to think of morality in terms of striving to be a member of a kingdom of ends, understood as an ideal community, in the early 1760s, and that he was influenced in this by his encounter with the Swedish mystic Emanuel Swedenborg. Swedenborg wrote volumes about his visions of heaven and hell, and in 1766 Kant published a book on Swedenborg, *Dreams of a Spirit Seer*, a commentary on Swedenborg's magnum opus, *Heavenly Secrets*. Most commentators take Kant's attitude towards Swedenborg to have been entirely negative, and argue that, at the most, Kant's encounter with him had a purely negative impact upon his development, inducing him to reject certain of his early metaphysical positions. I argue, however, that Swedenborg had a positive influence upon Kant's development, particularly upon his ethics, for Kant's conception of a kingdom of ends is modeled upon Swedenborg's conception of heaven as a community of spirits governed by moral laws. In other words, I argue that Kant's idea of a kingdom of end is the idea of a spiritual community, and that reading Swedenborg had a major impact on the development of Kant's ethics because it provoked him to conceive of ethics in terms of striving to be a member of an ideal spiritual (or later in his development: "intelligible") world.

Although Kant's ethics may have been provoked by his reading of Swedenborg, his mature conception of a kingdom of ends is a radical advance on the simple idea of a community of spirits Kant was toying with in the early 1760s. In particular, the idea of a kingdom of ends is the idea of community of *autonomous* individuals. In **chapter four** I

examine the development of this idea of an ideal moral/spiritual community in more detail. In the early 1760s Kant was thinking of ethics in terms of striving to be a member of an ideal spiritual community, and he was conceiving of such a community as a community of individuals governed by pneumatic (spiritual) laws. At this stage in his development, however, he did not seem to have thought it to be important that these laws must be thought of as given by the members of the community. The mature Kant, in contrast, conceives of such a community as governed by moral laws and believes that the only type of individuals that can really be members of a community are autonomous agents. The reason for this is that an individual can only be a member of a unified (intelligible) world if the individual itself is the ground (or more precisely if it is the concurrent ground) of the intelligible relations between individuals that constitute the world, because for a world to be *a* world it must be unified and the principle (or source) of its unity must be intrinsic to the world. And this is only possible if the individuals that constitute the world are the source of the unity of the world. Now, what unifies a world, and makes the world *a* world, are the relations (or laws) that hold between its individual members. So for a world really to be *a* world the individual members of the world must be the source of the laws that provide the world with its unity. Thus the idea of a member of *a* world is the idea of a being who “gives”, or legislates, the laws of the world. As a consequence, the only type of being that has the capacity to be (or become) a member of a ‘world’, strictly speaking, is a being that is aware of potential laws and chooses to actualize these laws, and this is, by definition, an autonomous agent.

PART TWO - METAPHYSICS

In **chapter 5** I examine the place of the idea of a world in Kant's theoretical philosophy. I explain the status of rational cosmology in both his pre-critical and critical theoretical philosophy. Kant is famous for claiming that we can know nothing about the intelligible world. When Kant claims that we can know nothing of the intelligible world he is not claiming that we can say nothing interesting or useful about our *idea* of an intelligible world. Instead of banishing all metaphysical speculation as meaningless, he merely wishes to clarify the epistemic import of such speculation. Such speculation cannot provide us with any knowledge. The reason for this is that pure thought, governed by the law of non-contradiction, provides us with no criterion of real possibility. The fact that a concept does not contain a contradiction merely implies that the concept is thinkable. The fact that a concept is thinkable, however, does not imply that there actually is, or even could be, an object corresponding to the concept. Thinkability, then, is not an adequate criterion for real possibility. This commitment is the basis for Kant's radical break with the rationalist tradition. For if pure thought provides us with no criteria for real possibility we must posit some other, distinct, faculty (the faculty of intuition) that can explain how we are able to make such judgments.

In **chapter six** deals with Kant's theory of **interaction**. I place Kant's account of interaction in its historical context. By the time Kant began his philosophical career, there were three standard theories of interaction: pre-established harmony, occasionalism and physical influx or influence. There are various ways of characterizing the difference

between these three positions. The simplest is to explain it in terms of a finite substance's responsibility for (internal and external) change. The theory of *(physical) influence* asserts that individuals can cause changes both in themselves and in others – that is, they can cause both internal and external change. The theory of *occasionalism* denies that finite substances are the cause of change either in themselves or in others. The theory of *pre-established harmony* asserts that finite substances are the cause of changes in themselves, but not in others. In addition, it should be noted that in the early eighteenth century the dominant account of 'physical influx' involved (as the name suggests) the idea of the accidents of one substance 'flowing into' another substance. Following Kant I refer to this position as the theory of *crude physical influence*. I argue that Kant, although he rejects *crude physical influx*, can be understood as advocating a version of physical influx.

Chapter seven examines Kant's theory of **action**. Kant has a problem conceptualizing action because, like Leibniz, he is committed to what I call the Principle of Active Inherence. According to this principle, an accident (or what Kant refers to as a 'determination') can only truly inhere in or belong to a substance if the substance is the active cause or ground of the accident. A determination, then, is only the determination of a particular individual if the individual is somehow the 'ground' of the determination. I have named this doctrine the Principle of Active Inherence. If we accept the principle of active inherence, though, it is not clear how one individual can ever be the cause of any change in another individual. If a determination can only be a determination of individual *b* if *b* is the active ground or cause of the determination, how can another substance ever

be the cause of a change in *b*? Leibniz's solution was to admit defeat and conclude that one substance cannot be the cause of a change in another. Kant's solution to this problem will be to claim that we can understand the idea of an individual being acted upon, without appealing to the untenable notion of accidents flowing into the individual, in terms of the agent "determining the active power of the substance being acted upon" . This account of action does not violate the Principle of Active Inherence, because the patient's determination inheres in the patient because it is a result of the patient's power. This power, however, has been determined by the agent. The model Kant introduces to clarify the notion of one individual determining the power of another is that of the withdrawal of resistance. Individuals, on this model, already resist one another. And one individual substance (the agent) is the 'cause' of a change in another individual substance (the patient) if the change in the patient is the result of the agent withdrawing its *resistance*. The patient remains, however, essentially active, for the determination is the result of its power. Thus each individual is essentially active in that everything that happens to a particular individual (everything a particular individual suffers) is the result of its own power or potentiality. On this model of action, however, individuals can only act upon one another if they are *already* resisting one another, and Kant argues that individual can only resist one another if they are members of a *community*.

In **chapter eight** I examine the logical basis of Kant's idea of **community** in greater detail. The chapter is divided into three sections. In (*8a*), I examine his account of **the concept of community**, introduced in the table of categories in the *Critique of Pure Reason*. I explain how this category is related to the disjunctive form of judgment and

argue that **the category of community is the concept of a whole the parts (or members) of which mutually exclude one another**. In addition I explain what Kant means in claiming that the concept of resistance is a predicable of the category of community. This claim, that Kant makes in passing, is often overlooked by commentators. However, given the role the concept of resistance plays in his model of action examined in chapter seven, this claim is highly significant, for it explains why and how Kant believes that action (and interaction) is intelligible. In *(8b)*, I explain how **the idea of community** is to be distinguished from the *concept*. I begin by examining his general account of the distinction between concepts and ideas and I argue that the idea of community is distinguishable from the category of community in, at least, three ways: **in the case of the idea of a community: (i) its parts/members must be logically prior to the whole, (ii) its parts/members must be simple (i.e. they cannot themselves have parts) and, (iii) it must be an absolute whole, in the sense of being a whole that is not itself part of any other whole**. In *(8c)*, I explain Kant's distinction between the idea of an **ideal** community and that of a **real** community. Our idea of a **real** community is **the idea of a real as opposed to an ideal whole**. This distinction has to do with the nature of the *unity* of the whole. An ideal whole is a whole whose unity exists merely in the mind of the observer, say God. A real whole, on the other hand, is a whole whose unity is *intrinsic* to the whole, in the sense that the whole itself is the source of its unity. For Kant, a community is unified by inter-substantial laws, where laws are thought of as intelligible relations that bind the members of the community together. In the case of an ideal community there will be a harmony between the change of state of one substance and that of another. An ideal observer could recognize regularities between the change of

state of one substance and that of another and could formulate inter-substantial laws to capture these regularities. The individual substances themselves, however, are not governed by these laws, nor are they the source of the laws, for the laws only exist in the mind of the observer, and so although the observer experiences the individuals as a whole they are not, in themselves, a whole as there is nothing that unifies them. In a real community, on the other hand, the community itself must be the source of the laws. Now, as the community just is its members, then the members of the community themselves must be the source of the laws that provide the community with its unity. This is why a real community must consist of autonomous agents. For an autonomous agent is, by definition, an individual that “gives” (i.e. “is the source of”) laws.

In the final chapter, **chapter nine**, I demonstrate how understanding Kant’s theoretical account of community and interaction can help us understand his practical philosophy. I do this by showing how Kant’s theoretical understanding of community and interaction underlies his theory of property.

I have claimed that Kant argues: (a) An individual *a* acts upon another individual *b* if *a* withdraws some impediments which allows a change to occur in *b*. According to this model of action, the patient is the ground of both determinations, but the agent is the ground of the change of determinations. (b) Real interaction is only possible between members of a real community. And, (c) a real community is only possible if each individual member of the community has ‘given’ the laws that unify the community. This account of interaction is highly abstract. In chapter eight I offer a concrete illustration.

Kant's account of property in the *Doctrine of Right* of the *Metaphysical of Morals* is based upon an analysis of the ideal of “the civil condition” and I will show that Kant conceives of the ideal of a “civil condition” as community understood in these terms, and that his account of property (and in particular his account of the transferal of property) has to be read in the context of his metaphysical analysis of interaction and the idea of community.

Recognizing that Kant conceives of the ideal of a civil condition as the idea of a real community, derived from the category of community, helps us understand Kant's account interactions between individuals involving property. Individuals can interact in two ways with regard to property: they can assert rights against one another, and property rights can be transferred from one individual to another. In this chapter I examine both kinds of interaction involving property, paying particular attention to Kant's account of the transferal of property.

In this concluding chapter I attempt to bring together a number of claims made in the two halves of this dissertation. In the first part of the dissertation I argued that the idea of a good will or autonomous individual is the idea of a member of an ideal community, and that to be virtuous is to strive to be such an individual. In the second part I argued that for Kant the pure idea of a community is the idea of a real whole the members (or parts) of which are simple and logically prior to the whole and which mutually limit or resist one another. I also argued that such resistance is necessary for interaction between individuals. In the concluding chapter I argue that Kant believes that our idea of a

political community governed by juridical laws (or what Kant often calls the “civil condition”) is the idea of a community in this sense. Indeed, I suggest that Kant believes that the **only** way we can conceive of a real community is as a political community. If I am right then Kant believes that the idea of a kingdom of ends is the idea of an ideal political community, and that to be virtuous is to strive to be a member of such an ideal political community. Kant’s moral ideal, then, is the idea of an ideal political community.

If my interpretation is correct it implies a radical re-evaluation of Kant’s ethics, and in particular a reevaluation of the relationship between the *Doctrine of Right* and the *Doctrine of Virtue*. On the standard reading of Kant's moral philosophy his ethics is more basic than his ‘political philosophy’. Thus Rosen (1993) remarks that, “for a long time it was fashionable to regard Kant's political writings as minor works” (p.1). If I am correct, however, Kant's political philosophy stands at the heart of his ethics because the idea of a kingdom of ends is, and can only be, conceived of as a political community governed by juridical laws. For this reason Kant's ethics is based upon his so called ‘political philosophy’, for his doctrine of rights provides his ethics with its content. This is why the *Doctrine of Right* precedes the *Doctrine of Virtue*, for to be virtuous is to strive to be a member of an ideal political community governed by laws of right.

Part One

ETHICS

Chapter One

On the Purpose and Value of Moral Philosophy

Introduction

Many scholars working on Kant's ethics have been influenced by Rawls's claim that Kant should be read as a **moral constructivist**. According to this interpretation, our moral ideas, such as the idea of a kingdom of ends, are constructed by practical reason, which is understood to be a faculty entirely distinct from theoretical reason. Practical reasoning is governed by what Rawls and his followers call the Categorical Imperative (CI) Procedure, and our moral ideas are somehow constructed using this procedure. On such an interpretation the categorical imperative (procedure) is logically prior to the idea of a kingdom of ends, and we can investigate the nature of this procedure without engaging with Kant's theoretical philosophy. Rawls' interpretation has been extremely influential, and many of his former students are now prominent Kant scholars. One result of this influence has been the practice of studying Kant's ethics divorced from his theoretical work. I believe that such a divorce has led to an impoverished understanding of Kant's ethics.

I reject the claim that Kant is a moral constructivist, instead I maintain that he should be understood as a **moral idealist**. I argue, in contrast to Rawls and his followers, that our

moral ideas are theoretical ideas, and that they are independent of, and logically prior to, the categorical imperative.² In particular I argue that the idea of a kingdom of ends is a particular theoretical idea, namely the pure idea of a community, which can be traced back to the category of community introduced in the table of categories of the *Critique of Pure Reason*.

As a result of this, the idea of a *member* of a kingdom of ends is also a theoretical idea; it is the idea of a member of a community. This theoretical *idea*, however, presents itself to us practically as an *ideal*, that is as something of extreme value that we should strive to instantiate. The categorical imperative is merely the practical recognition by our conscience of the theoretical idea (of a member of an ideal community) as an ideal. For the categorical imperative commands us to strive to be a member of a kingdom of ends. Kant is a moral idealist and, in contrast to the moral constructivist, believes that our moral ideas are logically prior to the categorical imperative, for the command “be a member of a kingdom of ends!” presupposed that we understand the idea of a kingdom of ends.

The moral ideal, then, is the idea of a member of a kingdom of ends. The *content* of this ideal is a theoretical idea that can be understood and analyzed theoretically. The *value* that we place on this idea, however, is purely practical, for however much we analyze theoretically the pure idea of a community we will never discover the value that our conscience places upon the idea of being a member of a community. However, although

² And would be logically prior to the CI Procedure if Kant believed there was such a thing.

the value our conscience places upon the idea of being a member of a kingdom of ends is beyond any understanding, the idea of a kingdom of ends is something that can be fully understood theoretically. The bulk of this dissertation will be concerned with justifying my interpretation of Kant as a moral idealist, and with explaining Kant's analysis of what is involved in the pure idea of a community or kingdom of ends. In addition, I will examine the genesis of Kant's belief that morality consists in striving to be a member of an ideal community. I will argue that this can be traced back to Kant's engagement with Swedenborg, the Swedish spirit-seer, in the early 1760s, for Kant's idea of a kingdom of ends is modeled upon Swedenborg's vision of heaven as a community of spirits in interaction.

There is a strong tendency amongst contemporary philosophers to regard ethics as an autonomous discipline and what lies behind this tendency is the belief that one can, and should, do ethics without engaging with traditional theoretical or metaphysical questions. Such beliefs and tendencies are also prevalent in the field of Kant studies, and many Kant Scholars believe that we can do Kantian ethics without reference to his theoretical works. One of the primary goals of arguing that Kant should be understood as a moral idealist is to combat this tendency.

I maintain, then, that, for Kant to be virtuous is to strive to be a member of an ideal community, and I will argue that Kant believes that if we examine our pure idea of a community we will discover that the only way we can conceive of a *member* of a community is as an autonomous individual. This is why, for Kant, to be moral (or

virtuous) is to strive to be an autonomous individual. We possess, he believes, the pure idea of what it is to be such an individual, and to be virtuous is to make this idea the object of our will, or what Kant calls our faculty of desire. The fact that Kant believes that to be moral involves making a particular pure idea the object of our faculty of desire, suggests a role for traditional metaphysical inquiry beyond that of mere conceptual game playing. For, insofar as a pure idea is to serve as a guide for our conduct it is important that we grasp the idea clearly. In the first few sections of this chapter I will attempt to justify my claim that, for Kant, the idea of a good will is a theoretical idea, logically independent and prior to the categorical imperative, and that Kant's primary task in the *Groundwork* is to analyze this idea.

It is not immediately apparent what value such a theoretical analysis of or moral ideas actually has, for Kant, in contrast to intellectualists such as Wolff and Baumgarten, believes that we do not need to have a clear understanding of the moral ideal in order to act ethically. He believes that even the pre-philosophical individual, with an obscure (theoretical) understanding is perfectly capable of acting morally. A clearer theoretical understanding of the idea of a member of a kingdom of ends does not effect how this idea appears to our conscience as an ideal. In the final section of this chapter I will attempt to explain why Kant believes that theoretical clarification of the idea of a good will is of practical significance.

Kant wishes to steer a middle path between intellectualism and voluntarism. The intellectualist believes that the good is irresistible and that we always act under the aspect

of the good. As a consequence of this he believes that all immoral behavior is a result of a false (theoretical) understanding of what constitutes the good. The voluntarist, in contrast believes that we can freely choose either the good or the bad. Kant agrees with the intellectualist that the good is irresistible, but he also agrees with the voluntarist that choosing the bad is the result of free choice and not faulty understanding, for we can (theoretically) understand the right thing to do and yet freely choose not to do it. On the face of it, these two commitments seem incompatible. Kant's answer is that although the good is irresistible, in so far as we pay attention to it, we are free to not pay attention. To understand what Kant could mean by this we must examine his account of conscience. For Kant, the conscience is like a court and the voice (or judge) of conscience is our practical awareness of the value of the (theoretical) idea of being a member of a kingdom of ends. The voice of conscience, then, commands us to strive to be a member of an (ideal) kingdom of ends. This voice is like that of a siren, for, insofar as we listen to it, it is irresistible. Our freedom (in the negative sense) consists in the fact that we can refuse to listen to this voice. We can try to plug our ears, or to drown out its call by focusing our attention on other competing voices. For the court of conscience resembles a court, not only by having a judge, but also in the fact that there is a defense attorney. The defense attorney in the court of conscience is, Kant believes, like an internal voice that competes with the voice of the judge for our attention in the court of conscience, and it does this by constantly offering excuses. Immorality is, Kant believes, the result of freely choosing to listen to the voice of the defense attorney rather than the voice of the judge.

This model of the conscience helps us understand the value of moral philosophy. The purpose of moral philosophy is to clarify our moral ideas theoretically. Such theoretical clarification is not necessary for morality; we do not need to do philosophy, or listen to moral philosophers, in order to be virtuous. To be virtuous, all we need to do is pay attention to the voice of conscience. Engaging in moral philosophy, however, can make such attentiveness easier. For Kant believes that (theoretical) clarification of the idea of a good will, “makes it shine forth more brilliantly”; such clarification makes the idea more noticeable and easier to pay attention to. Such clarification, in effect, amplifies the voice of the judge in the court of conscience, and makes it harder for us to be distracted by competing voices. In addition, and perhaps more importantly, such conceptual clarification can help us quiet the voice of the defense attorney. This voice, Kant believes, offers excuses and these excuses are the product of theoretical reasoning. For example, the voice of the defense attorney tells us that it is either physically or logically impossible (for us) to be truly autonomous. A clearer theoretical understanding of the idea of an autonomous individual makes it harder for us to give ourselves excuses for not living up to this ideal; engaging in such clarification quiets the voice of the defense attorney, and in so doing makes it easier for us to listen to the voice of the judge.

In his pre-critical period, Kant was attempting to develop a monadology, but, unlike Leibniz, one in which there was real interaction. By the mid 1770s he had come to realize that such metaphysical speculation could not provide us with knowledge or cognition. He did not, however, abandon such speculation, for although such metaphysical speculation

cannot provide us with a picture of the way the world is in itself, it can provide us with an ‘image’ of what we can become.

I maintain, then, that Kant should be understood to be a **moral idealist**. On this interpretation, the idea of a kingdom of ends is a theoretical idea, namely the idea of a community of individuals in real interaction, and this idea is derived from the category of community, the third category of relation of the table of categories of the *Critique of Pure Reason*. An analysis of the idea of community reveals that the only way we can conceive of a member of a community is as an autonomous individual, for each individual member of a community must be the source of the laws that provide the community with its unity. This (theoretical) idea of being an individual member of a community (or, what amounts to the same thing, an autonomous individual) presents itself to our conscience as an *ideal*, as something of supreme value that we should strive to instantiate. Although the content of the idea of a kingdom of ends is purely theoretical, the *value* we place on the idea of being a member of such a community is purely practical, and no amount of theoretical analysis will ever reveal the importance our conscience places on striving to be a potential member of a kingdom of ends. Its value is, quite literally, unintelligible.

In this chapter I shall focus upon explaining and justifying two claims: firstly, that Kant should be understood as a moral idealist and, secondly, that, for Kant, the primary function of moral philosophy is to arrive at a clearer theoretical understanding of what it is we are striving to be insofar as we recognize the demands of morality.

In **section one** I compare my approach with that of Rawls and his followers who argue that Kant should be understood to be a *moral constructivist*. The moral constructivists argue that our moral ideas, including the idea of a kingdom of ends, are practical (as opposed to theoretical) ideas that are constructed on the basis of a practically rational procedure. In **section two** I explain my interpretation of Kant's strategy in the *Groundwork*. In opposition to the moral constructivists, who argue that Kant's strategy in the *Groundwork* is to discover the implicit procedure of practical rationality (the "CI-procedure") and then show how the idea of a kingdom of ends can be constructed by means of this procedure, I argue that Kant's goal in the *Groundwork* is to arrive at a clearer theoretical understanding of what it is that our conscience demands that we become, and that his strategy is to begin by examining our pre-philosophical moral intuitions to show that what the person who listens to their conscience is striving to be is a member of an ideal community. Such a person does not necessarily understand theoretically that this is what they are doing, and Kant's goal in the *Groundwork* is to provide such a person with a clearer theoretical understanding of what they are attempting to be when they listen to the judgment of their conscience. In **section three** I explain Kant's account of the relationship between the notions of virtue and holiness. I argue that Kant believes that it is our duty to be holy, and that to be virtuous is, by definition, to strive to be holy. Given Kant's belief that ought implies can, the fact that we have a duty to be holy implies that it is possible for us to be holy. I explain how the postulates of practical reason introduced in the *Critique of Practical Reason* are intended to explain this possibility. In **section four** I examine Kant's disagreements with Wolff and Baumgarten, two of his rationalist predecessors who advocated a rationalist and

perfectionist morality. I argue that Kant himself should be understood as being both a rationalist and a perfectionist, and that what he objects to is their *formalistic* formulation of rationalism and perfectionism, and I suggest that he would reject moral constructivism on similar grounds. In **section five** I examine a number of passages in the *Critique of Practical Reason* that seem to offer support for a moral constructivist reading of Kant. In particular Kant, famously, insists that it is “the moral law that first determines and makes possible the concept of the Good” (5:64). I argue that these passages do not have the implications they seem to have at first sight, for after making such remarks, Kant makes it clear that this priority is merely methodological and not ontological, for he believes that although the categorical imperative is the *ratio cognoscendi* of the idea of being a member of a kingdom of ends, the idea of being a member of a kingdom of ends is the *ratio essendi* of the categorical imperative.

In **section six**, the final section of the chapter, I examine Kant’s account of the *value* of engaging in moral philosophy. Kant is not an intellectualist, for he believes that to be virtuous all we have to do is listen to our conscience, so a clearer theoretical understanding of what our conscience demands of us is not necessary in order to be moral. Given his rejection of intellectualism, however, it is not immediately clear why arriving at a clearer theoretical understanding of what morality demands has any practical value. I argue that the theoretical clarification that moral philosophy provides has motivational benefits. The reason for this is because immorality is always the result of choosing not to listen to the judgment of our conscience, and this happens, Kant believes, when we listen to excuses for not living up to the moral ideal, and these excuses are

themselves the product of theoretical speculation. A clearer theoretical understanding of the moral ideal will not, in itself, make us more moral, but it will make it harder for us to give ourselves excuses for not living up to the ideal. Engaging is moral philosophy, then, quietens the excuse giving voice, the voice of the defense attorney in the court of conscience, which makes it easier for us to be attentive to the voice of the judgement of conscience.

(1a) Moral Constructivism & Moral Idealism

In his influential paper, “Themes in Kant's Moral Philosophy”, Rawls (1993) argues that, for Kant “the basic moral concepts of the right and the good, and the moral worth of persons, are not analyzable in terms of non-moral concepts” (p.303). According, to Rawls, then, our basic moral concepts are not theoretical concepts. Instead he argues that Kant should be understood to be a **moral constructivist**. According to this interpretation, moral concepts, including the idea of a member of a kingdom of ends, are not theoretical ideas, but practical ideas that are constructed by practical reason. “Practical reason” Rawls (2000) argues, “constructs for the will its own object out of itself” (p.230). To reason practically is to implicitly follow a certain procedure, and Rawls calls this procedure the Categorical Imperative (CI) Procedure, and he maintains that (at least some of) our moral ideas are produced by means of this procedure.

Although Rawls is famous for advocating moral constructivism, he has indicated, especially in his more recent work, that he believes that certain fundamental moral ideas

are not constructed. For example, in *Political Liberalism* he argues that “the conceptions of society and person as ideas of reason are not, certainly, constructed any more than the principles of practical reason are constructed” (p.108). A number of his followers, most notably Korsgaard, however, have explicitly argued for moral constructivism “all the way down”, believing that *all* of our moral ideas are constructed by practical reason. Although Rawls himself places limits on his own constructivism, he himself is at the very least committed to the view that, for Kant, the idea of a kingdom of ends is constructed. Thus, he argues in his *Lectures on the History of Moral Philosophy* (2000) that,

The realm of ends is an (ideal) object – a social world – the moral constitution and regulation of which is specified by the totality of precepts that meet the test of the CI-procedure (when these precepts are adjusted and coordinated by the requirement of complete determination). (p.225)³

According to the moral constructivists, then, the idea of a kingdom of ends cannot be defined independently of the CI-procedure. The notion of a moral world is, by definition, the idea of a world that satisfies the requirements of the CI-procedure, and this formal procedure is logically and definitionally prior to the idea of a kingdom of ends. Accordingly, Kant's primary aim in the *Groundwork* is to clarify what is involved in this procedure. When I criticize moral constructivist interpretations of Kant in this paper, my objections are primarily directed towards the “all the way down” constructivism of some of his followers, such as Korsgaard, and my arguments in this paper do not rule out the possibility that Kant believes that some of our, less central, moral concepts may be

³ Or, as Rawls (1993) explains it in ‘Themes in Kant’s Moral Philosophy’, the realm of ends “is simply the social world that would come about (at least under favorable conditions) if everyone were to follow the totality of precepts that result from the correct application of the CI-procedure” (p.301).

constructed.⁴ I do, however, reject Rawls' claim that our idea of a kingdom of ends is, itself, constructed.⁵

In claiming that Kant should be understood as a **moral idealist** I argue, in opposition to the moral constructivists, that our fundamental moral ideas, including the idea of a kingdom of ends, are independent of, and logically prior to, the notion of the categorical imperative or that of a CI-procedure. The categorical imperative demands that we strive to be a member of a kingdom of ends, and so the notion of the categorical imperative is definitionally dependent upon the idea of a kingdom of ends. Any theoretical understanding of the categorical imperative presupposes an understanding of the idea of a kingdom of ends.

One of the reasons for rejecting moral constructivism, especially the "all the way down constructivism" of Korsgaard, is that such a position makes it difficult to explain the unity of theoretical and practical reason. For, according to the moral constructivist,

⁴ Although, ultimately, I am unsympathetic to any interpretation of Kant that suggests that moral ideas are constructed procedurally. The reason for this is because I believe that there is no such thing as the CI procedure. Instead of identifying a certain *procedure*, the first formulation of the categorical imperative identifies a certain *capacity*. As beings with what Kant calls a predisposition to morality, we must possess the capacity to "give" universal laws. On this interpretation, to be practically rational (that is, to obey the categorical imperative) is not to follow a certain procedure (the CI procedure), but to realize a certain capacity (the capacity to "give" universal laws).

⁵ As we have seen, Rawls argues in *Political Liberalism* that the idea of "society" is not morally constructed, however he argues in his Lectures that the idea of a "realm [kingdom] of ends" is constructed. I agree with Rawls that our idea of society is not constructed, however I believe that, for Kant, the idea of a kingdom of ends just *is* the *a priori* theoretical idea of society, or what he calls community. Onora O'Neill (2003), argues that what distinguishes Rawls' account of justice from Kant's is that, "Kant's public is not the Rawlsian public, consisting only of fellow citizens in a bounded, liberal democratic society: it is unrestricted. Hence, Kant's conception of ethical method takes a cosmopolitan rather than an implicitly statist view of the scope of the ethical concern" (p.362). I suggest that the basis of this disagreement is that Kant identifies the idea of a kingdom of ends with our (unconstructed) theoretical idea of society (in general) whereas Rawls distinguishes between these two ideas

practical and theoretical reason are two distinct faculties governed by distinct procedures. At the very least the moral constructivist needs to give some account to explain why we are justified in calling both of these faculties reason. On my interpretation, in contrast, there is only one faculty of reason, the faculty Kant discusses in the *Transcendental Dialectic* of the *Critique of Pure Reason*. The objects (or products) of this faculty are the ideas of pure reason, and these ideas can have either a theoretical or a practical application. They can be applied theoretically to our faculty of intuition, and practically to our faculty of volition, or what Kant himself calls the faculty of desire or will. Both the human faculty of intuition and the human faculty of volition have a certain subjective form. The form of our faculty of intuition is space/time. The subjective form of our faculty of desire is act upon maxims. Theoretically, the ideas of pure reason can have only a *regulative* use, for no objects can be given in intuition, in space or time, corresponding to these ideas. Practically, however, these ideas can have a *constitutive* application, for ideas of pure reason, although they cannot be objects of intuition, can be objects of volition, for an idea of pure reason can be the object of a maxim. In particular, I can make it my fundamental maxim to be a member of a community.. On my interpretation, then, practical rationality is not, as the moral constructivists argue, to be defined procedurally; rather, one is practically rational if one chooses to make an idea of pure reason the object of one's volition. This is what one does if one listens to one's conscience and chooses to be a member of a kingdom of ends.

Kant frequently stresses his moral idealism. For example, in lectures on ethics from 1785, a lecture course given around the time of the publication of the *Groundwork* he explains that,

to expound morality in its full purity is to set forth an Idea of practical reason. Such Ideas are not chimeras, for they constitute the guideline to which we must constantly approach. . . We have to possess a yardstick by which to estimate our moral worth, and know the degree to which we are faulty and deficient. . . An ideal is the representation of a single thing, in which we depict such an idea to ourself in *concreto*. All ideals are fictions. We attempt, in *concreto*, to envisage a being that is congruent with the idea. In the ideal we turn the ideas into a model. . . The ideal is a *prototypon* of morality. (*Ethik Mrongovius*, 29:604-5)⁶

Here Kant makes it clear that morality demands that we strive to instantiate an idea. The reason for this is that we need some yardstick by which to estimate our moral worth, and the only thing that can adequately fulfill this function is an idea that can serve as a model or prototype. In this passage, Kant refers to this idea as an idea of *practical* reason. I claim that what Kant here calls an idea of practical reason is, in fact, a theoretical idea that is derived from the table of categories. Textual support for such an interpretation is provided in the *Dialectic* of the *Critique of Practical Reason*.

In the *Dialectic*, Kant introduces the three ideas of practical reason: the “ideas of God, of an intelligible world (the kingdom of God), and of immortality” (5:137), and he makes it

⁶ In other lectures he argues that the moral ideal, the yardstick by which we estimate our moral worth, is the idea of an ideal man. And he explains that “we conceive of man first of all as an ideal, as he ought to be and can be, merely according to reason, and call this Idea *homo noumenon*; this being is thought of in relation to another, as though the later were restrained by him; this is man in the state of sensibility, who is called *homo phenomenon*. The latter is the person, and the former merely a personified idea; there, man is simply under the moral law, but here he is a phenomenon, affected by the feelings of pleasure and pain, and must be coerced by the noumenon into the performance of duty” (*Ethik Vigilantius*, 27:593).

quite clear that these practical ideas are the products of pure reason, being derived from the table of categories. As such it is clear that they should not be understood as the product of some distinct faculty of practical reason. For example, in the course of introducing these three ideas, Kant argues that,

Every use of reason with respect to an object requires pure concepts of the understanding (categories), without which no object can be thought. These can be applied for the theoretical use of reason, that is, for cognition of that kind, only insofar as intuition (which is always sensible) is also put under them, and therefore merely in order to represent by means of them an object of possible experience. But here [i.e. in the critique of **practical** reason] *ideas* of reason, which cannot be given in any experience at all, are what I would have to think by means of categories in order to cognize an object. Here, however, our concern with these ideas is not for the sake of theoretical cognition of their objects but only with whether they have objects at all. Pure practical reason provides this reality, and **theoretical reason has nothing further to do in this than merely to think those objects through categories**, and this, as we have elsewhere clearly shown, can be done quite well without needing intuition (whether sensible or supersensible) because **the categories have their seat and origin in the pure understanding solely as the faculty of thinking, independently of and prior to any intuition**, and they always signify only an object in general, in whatever way it may be given to us. (5:136 – my emphasis)

Here Kant makes it explicit that our moral ideas are thought by theoretical reason through the categories and are not the product of some distinct faculty of practical reason.⁷ In addition, I argue that the second of these ideas, the idea of “an intelligible world (the kingdom of God)” is to be identified with the idea of a kingdom of ends. Support for this

⁷ I agree, then, with Brandt (1995) that, “whatever metamorphoses [his] other doctrines undergo, Kant never doubts the categories and thus the table of judgments as the foundation of his system as a whole. . . . All critique, transcendental philosophy, and metaphysics (of morals and of nature) has its foundation in the table of judgments” (p.1). I will argue that the moral notion of an autonomous agent is the idea of a member of an intelligible world, and that Kant conceives of such a world as a community. Community is, Kant argues, the third category of relations, and so we should be able to explain our idea of autonomy in terms of this category and the corresponding (disjunctive) form of judgment. I will attempt to do this in chapter 8.

interpretation is provided by the fact that Kant parenthetically identifies the idea of an intelligible world with “the kingdom of God” [*dem Reiche Gottes*], a phrase that he normally treats as equivalent to “the kingdom of ends”.⁸

Elsewhere in the *Critique of Practical Reason* Kant identifies these three ideas as the ideas of God, *freedom*, and immortality. Kant, then, seems to regard the idea of freedom and the idea of the intelligible world as interchangeable. The reason for this is because Kant believes that our idea of an intelligible world is, by definition, the idea of a community⁹, and he is committed to the view that the idea of a member of a community can only be the idea of a free (i.e. autonomous) individual. A defense of this claim is beyond the scope of this paper. I mention it now, however, because there will be occasions in this paper when my interpretation rests upon assuming this identification.

(1b) Kant's Strategy in the *Groundwork*

There has been much debate amongst Kant scholars about the relationship between the three formulations of the categorical imperative introduced in *Groundwork* II. Rawls and his followers stress the first formulation, the formula of universal law. Other commentators, most notably Wood and Guyer, stress the second formulation, the formula of humanity. I argue that the third formulation is primary, that Kant believes that it is our

⁸ Elsewhere I offer a fuller defense of this claim, offering textual evidence from the Cosmology sections of his lectures on metaphysics.

⁹ For textual evidence for the claim that Kant identifies the idea of an intelligible world with that of a community, see: 28:657, 29:851-3, 28:196, 28: 581-2, 28:45, 29:868 & 29:1006-7. In these passages Kant distinguishes between our idea of a world and that of a mere multitude and argues that our idea of a world is the idea of a multitude of individuals in community with one another.

duty to be a member of a kingdom of ends, and that the virtuous person is striving to be a member of such a kingdom. However, although the virtuous person is striving to be a member of such a kingdom, she does not necessarily understand, theoretically, that this is what she is doing. The strategy of the *Groundwork* is to provide such an individual with the theoretical understanding of what they are trying to do. In terms of a theoretical understanding of what duty demands, the third formulation is primary. Phenomenologically, however, the first formulation is primary, for Kant believes that even the philosophically uneducated can recognize that they have a duty to only act upon maxims that can be universalized, although they do not understand (theoretically) that the reason this is important to them is because they recognize the value of being a member of an ideal community.

The following analogy might help to clarify my account of Kant's goal and strategy in the *Groundwork*: A professor issues the following command to one of his students: "Shane! Pay thirty-five dollars and sign here!" The student can follow this order without really understanding what he has been ordered to do. Later, however, perhaps after discussing the matter with his roommate, he might come to understand that he has been ordered to join the American Philosophical Association. The student is immediately aware that he has been ordered to sign a piece of paper, just as the common human understanding is immediately aware that she should only act on maxims that can be universalized. When the student realizes that he has been ordered to join the APA, he has a clearer theoretical understanding of what he has just been ordered to do. The virtuous individual who listens to their conscience, and who only acts upon maxims that can be universalized is, so to speak, applying to join a kingdom of ends, without necessarily understanding *theoretically* that this is what they are attempting to do. Kant's goal in the *Groundwork* is to provide the common human understanding with such a theoretical understanding.

Instead of attempting to discover which actions are right, Kant's primary aim in the *Groundwork*, then, is to explain what it is that a virtuous human being (that is, one motivated by duty) is striving to be. And he begins by examining the moral consciousness of the pre-philosophical individual. Such an individual recognizes that she should only act upon maxims that can be universalized. At the start of *Groundwork I*, Kant explains his objective. He writes that,

We have, then, to **explicate** (*entwickeln*) **the concept of a will** that is to be esteemed in itself and that is good apart from any other purpose, as it **already dwells in natural sound understanding** and needs not so much to be taught as only to be **clarified** – this concept that always takes first place in estimating the total worth of our actions and constitutes the condition of all the rest. In order to do so, we shall set before ourselves the concept of *duty*, which contains that of a good will though under certain subjective limitations and hindrances, which, however, far from concealing it and making it unrecognizable, rather bring it out by contrast and make it shine forth all the more brightly. (4:397 – my bolding)

Here Kant makes it clear that we already possess the concept of a good will and that this concept dwells in natural sound understanding, and that what Kant proposes to do in the *Groundwork* is to **explicate** and to **clarify** this concept. The conclusion he will reach at the end of *Groundwork II* is that the concept of a good will is the (theoretical) idea of a member of a community, and that this is equivalent to the idea of an autonomous individual. In claiming that the concept of a good will “already dwells in natural sound **understanding**” he means that the concept of a good will is a *theoretical* idea that can be understood theoretically. The concepts of duty “contains that of a good will” because it is our duty to be a good will, for when we understand the CI we understand that it is our duty to strive to be a member of an ideal kingdom of ends. The aim of the *Groundwork*,

then, cannot be to teach us this ideal, for it is an idea that we are already aware of, albeit obscurely. Instead, the aim of the *Groundwork* is to explicate and clarify this concept.

Kant is more explicit about this in the *Metaphysics of Morals*, where he writes that

no moral principle is based, as people sometimes suppose, on any *feeling* whatsoever. Any such principle is really an obscurely thought *metaphysics* that is inherent in every human being because of his rational predisposition, as a teacher will readily grant if he experiments in questioning his pupil socratically about the imperative of duty and its application to moral appraisal of his actions. **The way the teacher presents this (his technique) should not always be metaphysical nor his terms scholastic, unless he wants to train his pupil as a philosopher.** But his thought must go all the way back to the elements of metaphysics, without which no certitude or purity can be expected in the doctrine of virtue, nor indeed any moving force. (6:376)

To teach ethics is to clarify an obscure metaphysics, and to do it philosophically involves doing it in scholastic terms. Although Kant's ultimate aim in the *Groundwork* is to clarify the idea of a good will, he begins with an examination of the concept of duty, which will somehow help us clarify the concept of a good will. For Kant believes that the pre-philosophical consciousness of one's duty is a subjective, felt consciousness of the moral ideal. The consciousness of duty or obligation is a consciousness that has a certain idea as its object. Thus, if we want to clarify what is involved in this idea we must start by examining our consciousness of duty, for it is an (obscure) consciousness of the idea.

Our subjective consciousness of the categorical imperative, then, is the way the (theoretical) idea of a good will (that is, idea of an autonomous individual or a member of kingdom of ends) subjectively presents itself to our conscience. The categorical

imperative, then, is based upon a subjective and obscure consciousness of the pure idea of a good will. Kant's aim in examining this consciousness of obligation, then, is to 'clarify' what is involved in the pure idea of a good will. Our common moral judgments are always based upon an obscure representation of duty. Most people couldn't explain what is involved in the idea of duty, but they know how to act upon it. In *Groundwork* I & II Kant investigates our subjective consciousness of duty. The formulations of the categorical imperative are attempts to provide the reader with a clearer understanding of what is involved in the concept of duty. This clarification of what is involved in duty, however, is not, in itself, the aim of the *Groundwork*. Duty itself has an object, and this, ultimately, is what Kant is interested in. It is our duty to choose to be a particular type of being. Our obscure consciousness of duty involves an obscure consciousness of the type of being we have a duty to choose to be. To clarify our consciousness of the concept of duty is to clarify our concept of this being, until it shines forth as the idea of a member of the intelligible world.

What does it mean for a concept to be obscure?

Kant explains that his task in the *Groundwork* is to "explicate" and "clarify" the concept of a good will (4:397). To 'clarify' a concept is to move from an obscure to a clear understanding. 'Explication' is the (or at least a) way to clarify a concept, for to 'explicate' a concept is to make facts that are analytically true of the concept explicit.¹⁰

¹⁰ In his *Jäsche Logic* Kant distinguishes between two different types of identity in analytic judgments. The identity can either be 'explicit' or 'implicit'. If the identity is 'explicit' then the analytic judgment is tautological and as a result is 'empty' or 'fruitless'. For example, 'a dog is a dog' is tautological for the identity between the concept 'dog' and 'dog' is clear to everyone. "Propositions that are identical *implicite*,

That is, to explicate a concept is to become conscious of propositions which are analytically true of a concept but which we were previously not conscious of. Now, it is only possible not to be conscious of an analytic truth concerning a particular concept if we actually possess the concept but do not clearly see all the ‘marks’ of the concept. That is, for explication of a concept to be possible we must possess the concept, but our awareness of it must be **obscure**. In the process of explication then, an obscure concept becomes clearer. Kant explains the distinction between clarity and obscurity in the *Jäsche* Logic. He explains it in the following terms:

In every cognition we must distinguish matter, i.e. the object, and form, i.e. the way in which we cognize the object. If a savage sees a house from a distance, for example, with whose use he is not acquainted, he admittedly has before him in his representation the very same object as someone else who is acquainted with it determinately as a dwelling establishment for men. But as to the form, this cognition of one and the same object is different in the two. With one it is mere intuition, with the other it is intuition and concept at the same time. // The difference in the form of cognition rests on a condition that accompanies all cognition, on consciousness. If I am conscious of the representation, it is clear; if I am not conscious of it, obscure. (9:34)

The ‘savage’, then, sees something when he looks at the house, but his consciousness (of the form) of the house is obscure because he has no concept of a house. He intuits the matter of the house but does not understand its form. In claiming that our idea of a good will needs to be ‘clarified’ Kant suggests that our pre-philosophical consciousness of the idea of a good will is analogous to the ‘savage’s’ awareness of the house.

on the other hand, are not empty of consequences or fruitless, for they make clear the predicate that lay undeveloped (*implicite*) in the concept of the subject through explication (*explicatio*)” (9:111).

Kant has something like the following analogy in mind. In having the consciousness of duty we all, so to say, practically ‘intuit’ the pure idea of a good will, but this is not the same as to possess a clear theoretical understanding of the idea.¹¹ Kant begins the *Groundwork* from the perspective of innocent popular moral consciousness, and such a consciousness involves a consciousness of the pure idea of a good will in the same way that a ‘savage’ is conscious of a house. The ‘savage’ can recognize the house, and is able to identify and reidentify it, and could presumably find his way there, but he does not know what it is. Analogously the ‘moral savage’, or the common human understanding, can recognize the concept of a good will but does not really know what the idea of a good will is an idea of. The morality of the innocent common human understanding which bases its ethics on the *obscure consciousness* of duty and obligation, and which is Kant’s subject in *Groundwork* I, has, then, a similar relationship to the pure idea of a good will as the ‘savage’ has to the house. To provide a ‘savage’ with clarity would be to teach him the concept ‘house’. To bring clarity to the common human understanding, then, involves teaching such an understanding the pure concept of a good will. This involves changing the consciousness of the idea of a good will from an obscure consciousness into a true understanding.

(1c) Virtue and Holiness

Kant distinguishes between the ideas of virtue and holiness and argues that it is our duty to be holy. I argue that the idea (of a good will) that Kant wishes to ‘explicate’ in the

¹¹ As we have seen, I suggest that there are certain structural similarities between our faculty of intuition and our faculty of volition. Both have ‘objects’ and both have a certain subjective, human form.

Groundwork is the idea of a holy will, and his conclusion is that the idea of a holy will is the idea of an autonomous individual. To be virtuous is to strive to be such an individual.¹²

The idea of a holy will is the idea of a non-sensible and non-sensuous individual. As the idea of a non-sensible individual it is the idea of an intelligible individual, that is, it is the idea of an individual in so far as it is thought of as not a (possible) object of spatio-temporal intuition. It is an idea of a non-sensuous individual in the sense that it is the idea of a being that is not motivated by sensuous needs. It is the idea of a being that is not necessitated. For in thinking of such an ideal being, we are conceiving a being for whom morality is not a matter of demands, or imperatives, for such a being has no needs conflicting with laws of the community it is a member of; such a being wills these laws gladly.

In the *Groundwork* and elsewhere, Kant makes it clear that the idea of a good will, the practical ideal he wishes to analyze in the *Groundwork*, is the idea of a holy will and not that of a virtuous will. Thus, in the *Metaphysics of Morals*, he argues that,

Virtue so shines as an ideal that it seems, by human standards to eclipse holiness itself, which is never tempted to break the law. Nevertheless, this is an illusion arising from the fact that, having no way to measure the degree of a strength except by the magnitude of the obstacles it could overcome. . . we are led to mistake the subjective conditions by which we

¹² Kant famously begins *Groundwork I* with the claim that, “[i]t is impossible to think of anything at all in the world, or indeed even beyond it, that could be considered good without limitation except a good will” (4:393). It is my contention that the idea of a good will (i.e. the idea of an autonomous being or member of a kingdom of ends) is the idea of a holy will, and to be virtuous is to strive to be holy. One could say that the **moral constructivist** believes that virtue is logically prior to holiness, for we can define holiness as: “that what a virtuous individual strives to be”. The moral idealist believes that the idea of holiness is logically prior to virtue, for a virtuous individual is defined as: “an individual who strives to be holy”.

assess the magnitude for the objective conditions of the magnitude itself.
(6:397)

Similarly, in *Groundwork II*, Kant argues that,

A perfectly good will would. . . equally stand under objective laws (of the good), but it could not on this account be represented as *necessitated* to actions in conformity with law, since of itself, by its subjective constitution, it can be determined only through the representation of the good. Hence no imperatives hold for the divine will and in general for a holy will: the “ought” is out of place here, because volition is of itself necessarily in accord with the law. Therefore imperatives are only formulae expressing the relation of objective laws of volition in general to the subjective imperfection of the will of this or that rational being, for example, of the human will. (4:414)

Here Kant makes it quite clear that the idea of a good will is the idea of a holy will and not the idea of a virtuous human will that is “necessitated” and acts out of duty.¹³ The idea (of a good will), then, that Kant wishes to explicate in the *Groundwork* is not the idea of a (human) being that acts out of duty and obeys imperatives. In the *Groundwork*, Kant is attempting to provide the foundations for the analysis of the idea of such a holy individual, for he believes that the idea of such a holy will is an ideal which must serve, as we have seen, as a yardstick by which we can judge our own moral worth.

We can find passages in which Kant seems to give priority to the idea of virtue rather than that of holiness. For example, in the *Critique of Practical Reason* he argues that,

¹³ Further evidence for this is provided in his lectures from the early 1790s. In these lectures he argues that we have the idea of what it is to be a human being within us, and he argues that “humanity itself, if we wished to personify it, actually lacks any inclination to evil, but the more a man compares himself therewith, the more he finds out how far away he is from it” (*Ethik Vigilantius*, 27:609). As we have seen, the idea of a being that lacks any inclination to evil is a holy being. As we shall see in the following chapter, Kant often refers to the moral ideal as the idea of humanity personified. This passage should make it clear that our idea of ‘humanity personified’ is the idea of what he calls a holy being.

The moral level on which a human being (and, as far as we can see, every rational creature as well) stands is respect for the moral law. The disposition incumbent upon him to have in observing it is to do so from duty, not from voluntary liking nor even from an endeavor he undertakes unbidden, gladly and of his own accord; and his proper moral condition, in which he can always be, is *virtue*, that is, a moral disposition *in conflict*, and not *holiness* in the supposed *possession* of a complete *purity* of dispositions of the will. (5:84)

Here Kant seems to suggest that it is “incumbent upon us” to try to be virtuous and not holy. This seems to undermine my claim that the idea of a good will is the idea of a holy will. Here Kant seems to suggest that holiness is not even something we should strive for. However, this suggestion is uncharacteristic. Indeed, on the previous page Kant argued that

That law of all laws. . . like all moral precepts of the Gospel, presents the moral disposition in its complete perfection, in such a way that as an ideal of holiness it is not attainable by any creature but is yet the archetype which we should strive to approach and resemble in an uninterrupted but endless progress. (5:83)

Here Kant makes it clear that the moral ideal is the idea of being holy, which suggests that the moral law, or categorical imperative, commands us to strive towards holiness.¹⁴ Indeed, the moral law cannot command us: be virtuous! For, as Kant argues in the *Metaphysics of Morals*, “virtue itself, or possession of it, is not a duty (for then we would have to be put under obligation to duties)” (6:405). Instead, Kant believes that to be virtuous is to strive towards holiness and that another formulation of the categorical

¹⁴ And he adds that, “if a rational creature could reach the stage of thoroughly liking to fulfill all moral laws he would have achieved this ideal” (5:83).

imperative is: be holy! I believe that this is Kant's considered position. He makes it clear that this is his position in the ethics lectures he delivered at the time he was working on the *Groundwork*. In these lectures he proclaims that,

The ideal of the gospels has the greatest moral purity. The ancients had no greater moral perfection than that which could come from the nature of man, but since this was very defective, their moral laws were also defective. . . The principles of morality are [in Christianity] presented in their holiness, and now the command [i.e. Imperative] is: **You are to be holy**. (*Ethik Collins*, 27:252 – my emphasis)

Here Kant makes it clear that it is his belief that it is our duty to be holy.¹⁵ Similar passages are not hard to find.¹⁶

The duty to be holy, “ought implies can” and the postulates of practical reason

In the *Critique of Practical Reason* Kant introduces two postulates of practical reason. He argues that, for practical reasons, we must assume our immortality and the existence of God. The reason we must make these assumptions is because we have a duty to be holy, and Kant believes that ought implies can, and so this duty implies that it is not impossible for beings such as ourselves to achieve holiness. However, this is only possible if we assume that our duration in time is endless and that there exists an omnipotent, omniscient and just ruler of the world.

¹⁵ And in this passage he is clearly advocating what he calls “Christian” morality, for he makes it clear that he believes that the pre-Christian laws (or principles) were “defective”.

¹⁶ For example, in the same lectures he argues that “the principle we draw from the weakness of human nature is this: moral laws must never be laid down in accordance with human weakness, but are to be presented as holy, pure and morally perfect, be the nature of man what it may. . . the moral law is the archetype, the yardstick and the pattern of our actions [i.e. it is what Kant calls an ideal]. But the pattern must be exact and precise. . . The highest duty [of the moral philosopher?] is therefore to present the moral law [this ideal] in all its purity and holiness, just as the greatest crime is to subtract anything from its purity” (*Ethik Collins*, 27:294).

Kant, then, believes that it is our duty to be holy, for the moral ideal is the idea of a holy being. He also, famously, believes that ought implies can. Taken together, these two commitments seem to imply that Kant must believe that it is possible for us to become holy. There are good reasons, however, for thinking that it is impossible for beings such as we are to be holy. In this section I will briefly examine how Kant attempts to avoid this seeming contradiction. His response is to appeal to divine assistance. This is the function of the postulates of practical reason. Kant explains his general strategy in his lectures on ethics, explaining that,

The ideal of holiness. . . is the most perfect ideal for it is an ideal of the greatest purely moral perfection, but because such a thing is unattainable by man, it is based upon a belief in divine assistance. (*Ethik Collins*, 27:521)

Here, once again, Kant clearly identifies the moral ideal with the ideal of holiness, and he suggests that not only must we strive for holiness, but that we can legitimately hope to attain it, albeit with divine assistance. Our hope that it is possible to achieve holiness can be broken down into two components. To demonstrate that this hope is not irrational, Kant must demonstrate (a) that it is not contradictory to hope that we are capable of perfect virtue and (b) given the conceivability of perfect virtue, it is not contradictory to hope that, with divine assistance, we could achieve holiness.

(a) Given our temporal nature, it seems impossible for us to think of ourselves as perfectly virtuous, let alone as potentially holy beings. At any particular moment in time, the most we can say is that up until now we have acted out of duty, and indeed, we are all

aware that there have already been times when we have neglected our duty. At the most, then, we can think of ourselves as becoming more and more virtuous, but it is impossible for us to think of ourselves being perfectly virtuous, let alone holy. Once again, if ought implies can, it is difficult to see how Kant can consistently maintain that we ought to make it our aim to be holy, for he seems to believe that we cannot even hope to be perfectly virtuous. Kant's response to this problem once again involves an appeal to God, but this time not to his will (assistance) but to his intuition. For although we must experience our character as developing in time, as becoming more and more virtuous, we can hope that our existence in time is infinite and that God can experience what we can only experience as our never-ending increase of virtue as a perfectly virtuous disposition.

Thus Kant argues in the *Critique of Practical Reason* that,

the *eternal being*, to whom the temporal condition is nothing, sees in what is to us an endless series the whole of conformity with the moral law, and the holiness his command inflexibly requires in order to be commensurable with his justice in the share he determines for each in the highest good is to be found whole in a single intellectual intuition of the existence of rational beings. . . [The human being] cannot hope, either here or in any foreseeable future moment of his existence to be fully adequate to God's will. . . he can hope to be so only in the endlessness of his duration (which God alone can survey). (5:123)

Just as we can grasp that the series $1 + \frac{1}{2} + \frac{1}{4} + \text{etc.}$ converges on one, we can hope that God can grasp our ever increasing virtue as converging on perfect virtue. In so doing, Kant believes that we can reasonably hope that it is possible for us to be perfectly virtuous, at least in the eyes of God.

(b) This, however, brings us to a second problem. For even if Kant can convince us that there is nothing contradictory in hoping to be perfectly virtuous, such perfect virtue cannot be equated with holiness, for even if we are capable of always doing our duty it is not within our power to guarantee that we will always do it gladly, for we are beings with needs, and there are times when these needs conflict with morality. Even if we always do our duty, then, these needs make it impossible for us to do our duty “gladly” and with a “voluntary liking”.¹⁷ As we have seen, however, a holy being is good gladly. A holy being, then is both good and happy. It not only does what is good, but it is glad to be doing it. Kant's response is to argue that if we do our duty we can hope to be happy, but that this hope can only be based upon the hope of divine assistance, for we must hope that there is a God who assigns happiness in proportion to virtue. If this were the case then a perfectly virtuous individual would be perfectly happy, and would not merely obey the law, but do so gladly. This is Kant's doctrine of the highest good.¹⁸ As Kant explains this in the *Religion*:

If the strictest observance of the moral laws is to be thought of as the ushering in of the highest good (as end), then, since human capacity does not suffice to effect happiness in the world proportionate to the worthiness to be happy, an omnipotent moral being must be assumed as ruler of the world, under whose care this would come about. (6:8)

A human being who always acted from duty does not have a holy disposition if she, at times, does not do her duty gladly. This will be the case whenever our duty is in conflict

¹⁷ It should be noted that if we necessarily did our ‘duty’ gladly it would not be duty.

¹⁸ Although I believe that Rawls is wrong to suggest that our idea of a kingdom of ends is constructed practically, I believe a stronger case could be made for the claim that Kant believes that our idea of the “highest good” is constructed practically.

with our sensuous needs. Kant, however, believes that there is nothing contradictory in the idea of a being who brings it about that human beings are happy in proportion to their worthiness to be happy. An individual who is perfectly virtuous, that is, an individual who always does her duty, deserves perfect happiness. Such an individual, however, can only be perfectly happy if everything she does, she does gladly, so in hoping that there exists an omnipotent being who decrees that happiness be rewarded in proportion to virtue, we are hoping that the world is so organized so that a perfectly virtuous individual will always do her duty gladly. This is to hope that if we always do our duty all of our needs will (necessarily) be met. Thus, if it is possible for us to be perfectly virtuous, it is possible for us to believe that we can have a holy disposition. There is, Kant believes, nothing contradictory in this hope.

(1d) Kant's Disagreement with Baumgarten & Wolff

To understand Kant's ethical project, and to justify my interpretation, it is helpful to look at the contrasts he draws between his own position and that of Wolff and Baumgarten. Wolff and Baumgarten are (a) rationalists and (b) perfectionists. They believe that being moral is synonymous with being rational and that the principle of morals is to seek perfection wherever one can. Although Kant attacks their position, he himself can also legitimately be regarded as a moral rationalist and perfectionist. For he identifies the idea of a good will with that of a rational being and he believes that we should take holiness, i.e. the idea of something perfect, as our moral ideal. His main complaint against their rationalism and perfectionism is that it is merely formal and lacks content. Understanding

what Kant means by this and how he differentiates his own brand of rationalism and perfectionism from that of Wolff and Baumgarten will help us to understand his own position. It will also provide us with additional grounds for rejecting the moral constructivist interpretation of Kant.

(i) Rationalism

Kant agrees with Wolff and Baumgarten in their identification of morality with rationality. To be moral would be to be perfectly rational. Kant's criticism of Wolff and Baumgarten is that they do not give any account of what it is (or more accurately, what it would be) to be perfectly rational. This is why he accuses them of formalism. Many commentators find it ironic that Kant accuses others of formalism, because this is one of the most common criticism to be leveled against Kant's own ethics. Kant, however, believes that his moral theory is not 'formalistic' in the sense that Wolff and Baumgarten's is, for he believes that he can give a non-formal definition of what it is for a human being to be practically rational; for Kant a practically rational human being is one who strives to be a member of an ideal kingdom of ends. For Wolff, to have a rational will is merely to will in a certain way, namely rationally. For Kant, in contrast, practical rationality is defined in terms of the object of volition. A perfectly rational faculty of desire would be a faculty of desire that chooses as its object the idea of being an autonomous individual; and the idea of an autonomous individual is an idea of pure reason. Kant, then, criticizes Wolff for offering a purely formal definition of a rational will. He objects, then, to their purely formal characterization of the object of rational

desire. In contrast he believes that he can provide a material definition of what it is (or would be) to have a perfectly rational will.

The distinction between Kant's ethics and Wolff's is analogous to the distinction between transcendental and general logic

Towards the end of his Preface to the *Groundwork*, Kant compares his introduction to moral philosophy with that of Wolff, noting that,

Just because it was to be a universal practical philosophy it [Wolff's propaedeutic to moral philosophy] took into consideration, not a will of any special kind, such as one that would be completely determined from *a priori* principles without any empirical motives and that could be called a pure will, but rather volition generally, with all the actions and conditions that belong to it in this general sense; and by this it differs from a metaphysics of morals in the same way that **general logic**, which sets forth the actions and rules of *thinking in general*, differs from **transcendental philosophy**, which sets forth the special actions and rules of **pure** thinking, that is, of thinking by which objects are cognized completely *a priori*. For the metaphysics of morals has to examine the idea and the principles of a possible *pure* will and not the actions and conditions of human volition generally, which for the most part are drawn from psychology. (4:391)

Kant makes a similar point in his ethics lectures, lectures delivered at roughly the same time as he was writing the *Groundwork* (1785). Here Kant explains that,

Baumgarten and Wolff say that duty is the necessity of an action according to the greatest and most important grounds of motivation. Now to them it is all one, whether these grounds are from inclination or from reason. **It is thus no pure philosophy that they have in view here, but rather a general practical philosophy.** The latter treats of concepts and all actions that proceed from willing. How we ought to act it does not consider. It makes no mention of the determinations of our willing by pure motivating grounds of reason, but speaks in general of the determinations of the will. In general practical philosophy, nothing of morality must appear. (*Ethik Mrongovius*, 29:599 – my emphasis)

In both of these passages Kant suggests that the distinction between his own ethics and Wolff's is analogous to the distinction between transcendental and general logic he has introduced in the *Critique of Pure Reason*. In the *Introduction to the Critique of Pure Reason*, Kant explains that,

general logic abstracts from all contents of the cognition of the understanding and of the difference of its objects, and has to do with nothing but the mere form of thinking. (A54/B78)

In so far as Wolff's ethics is general, then, it abstracts from all contents of the faculty of desire and has to do with the mere form of willing. Transcendental logic, on the other hand, does not abstract from the content of thinking, but examines the *a priori* relationship between the faculty of thinking (the understanding) and the faculty through which objects are given to us (intuition). Thus Kant writes that, assuming there can be

concepts that may be related to objects *a priori*, not as pure or sensible intuitions but rather as acts of pure thinking, that are thus concepts but of neither empirical nor aesthetic origin [that is, that there are *pure* concepts], we provisionally formulate the idea of a science of pure understanding and of the pure cognition of reason, by means of which we think objects completely *a priori*. Such a science, which would determine the origin, the domain and the objective validity of such cognitions, would have to be called transcendental logic. . . (A57/B81)¹⁹

The difference, then, between general and transcendental logic is that whereas general logic is purely formal, transcendental logic is concerned not only with the form of thought but with (*a priori*) objects of thought. Thus Pozzo (1998) explains that,

¹⁹ Kant also explains this distinction in his lectures, claiming that transcendental logic "distinguishes itself from general pure logic in this, that the latter occupies itself with the mere form of the use of our understanding, but the former concerns the determination of the pure cognition of objects through the mere understanding. A pure concept of the understanding is a pure cognition of the object through the mere understanding, and the full opposite <*oppositum*> of the empirical concept, in that it is entirely thought purely *a priori*" (*Metaphysic Vigilantius*, 29:984).

“transcendental logic differs only slightly from general logic, but in its core; the former has no object at all, the latter has an object, a very general one, but always an object” (p.301-2).²⁰ As a consequence of this, Kant believes that general logic can only provide us with a negative criterion for truth, for truth does not have to do with the purely logical relationship between concepts, but has to do with the relationship between concepts and objects. The principle of general logic is the principle of non-contradiction, and Kant believes that “although a cognition may be in complete accord with [general] logical form, i.e., not contradict itself, yet it can still always contradict the object” (A59/B84). As a result of this Kant believes that principle of general logic can only provide us with a negative criterion of truth. What Kant means by this is that general logic can only tell us whether a concept (or proposition) is self-contradictory or not. Whether or not a concept is self-contradictory (or whether or not a proposition implies a contradiction), however, is not an adequate criterion of truth, for it may be the case that although proposition does not contain a contradiction, the proposition is not true because there cannot (in fact) be an object corresponding to the subject concept. In other words, although general logic can determine whether a concept is thinkable (which is to determine whether or not it contains a contradiction) it cannot tell us whether a thinkable concept has what Kant calls objective reality or validity, for inspecting a concept cannot tell us whether or not there could actually be an object corresponding to the concept. Transcendental logic, in

²⁰ Pozzo convincingly argues that Kant's distinction can be traced back to the traditional distinction between general and special logic, a distinction that can be traced back to Averroes' commentaries upon Aristotle's *Organon*. General logic examines the rules of thought in general, rules, Kant explains, “without which there can be no employment of the understanding” (A52/B76). Special logics, in contrast, examine the rules of thought about particular objects. Kant's distinction between general and transcendental logic is derived from this distinction, with transcendental logic being the most general special logic, concerned with the a priori relationship of thought (the understanding) to the faculty through which objects are given (intuition).

contrast, can be “a logic of truth” (A62/B87), for it has to do with the *a priori* relationship between concepts (the understanding) and objects in general. There can be a “logic of truth” because the faculty through which objects are given to us (intuition) has a certain form that we are aware of *a priori*. Thus, (the positive part of) transcendental logic examines *a priori* the relationship between the faculty of concepts (the understanding) and the faculty through which objects are given (intuition). Such a logic can provide us with a positive touchstone for truth because there are certain things we know *a priori* about what type of objects can be given in intuition, for our form of intuition is, Kant maintains, *a priori*. As a result of this, we can know *a priori* that an object can exist corresponding to certain concepts if the object is the sort of thing that can be given in space and/or time.

We can now return to Kant's distinction between Wolff's ethics and his own, a distinction that is analogous to the distinction between general and transcendental logic. A ‘general’ ethics will be one that is purely formal and does not concern the relation of the faculty of desire to any object. In addition, its principle will be a principle that provides us with a purely negative criterion for morality in the same way that the principle of non-contradiction (the principle of general logic) provides us with a purely negative criterion of truth. An ethics, like Kant's, that is analogous to transcendental logic will, in contrast, be an ethics that has some *a priori* content, in that it will examine *a priori* the relationship between the faculty of desire and the faculty through which objects are given to us. Transcendental *theoretical* philosophy examines the *a priori* relationship between the intellect and the faculty of intuition. Transcendental *practical* philosophy examines the *a*

priori relationship between the faculty of desire (volition) and the intellect. For, just as the faculty of intuition provides the intellect with its objects, the (faculty of) intellect can provide the faculty of desire (volition) with its objects, for ideas can be objects of volition. Such a “transcendental” (as opposed to a “general”) ethics, then, will examine the way in which ideas can be objects of volition, and this examination can be done *a priori*.

We are now in a better position to understand the distinction Kant draws between his ethics and Wolff’s. The difference between Kant’s way of proceeding and Wolff’s is that Kant thinks that the idea of a good will is a **pure** idea which can be defined purely from *a priori* principles and that its definition can be given purely in terms of *a priori* concepts and will not involve any reference to (or abstraction from) empirical motives or to any particular human capacity or faculty. The difference between this approach and Wolff’s is that Wolff believes that we can explain what it is to be a good faculty of desire merely in terms of explaining *how* such a faculty of desire wills, without explaining *what* it is that a good faculty of desire wills.

Another way to understand how Kant’s rationalism differs from Wolff’s is to examine Kant’s distinction between the idea of a perfectly rational human being (what I shall call a ‘being that has a perfectly rational faculty of desire’) and the idea of what he calls a rational being as such. A perfectly rational human being is one who always acts out of duty. The idea of a rational being as such, however, cannot be defined in terms of duty, because it is only a contingent fact about our nature that makes ethics a matter of duty for

us, and an account of what it is to be moral should be an account which is applicable to all conceivable (rational) beings, and not just to human beings. For to act out of duty is to obey an imperative and Kant believes that not all (conceivable) moral agents are bound by imperatives. Thus what it is to be moral cannot consist in acting out of duty. For, as Kant argues in the *Groundwork*,

no imperatives hold for the *divine* will and in general for a *holy* will: the 'ought' is out of place here, because volition is of itself necessarily in accord with the law. Therefore imperatives are only formulae expressing the relation of objective laws of volition in general to the subjective imperfection of the will of this or that being, for example, of the human will. (4:414)

According to Kant, then, our definition of a 'rational being as such' (or what he calls 'a good will' in the *Groundwork*) is logically prior to, and must be independent of, the definition of a 'being that has a rational faculty of volition', for Kant defines 'a being which has a rational faculty of desire' as 'a being that desires to be a rational being as such'. For us to have the capacity to be rational, then, we must possess the idea of what it is to be a rational being as such. We can only understand the notion of 'a perfectly rational faculty of desire' (that is, a good faculty of willing) by appealing to this pure idea of a good will. For a perfectly rational faculty of desire is a faculty of desire that has chosen to make the idea of a good will the object of its choice or desire. What it is to have a '*perfectly rational* faculty of desire', then, is defined in terms of the object such a faculty of desire takes.

A good (or rational) will, then, is not defined as a will (faculty of desire) that possesses “practical rationality” in the Rawlsian sense of being guided by the CI procedure. A good flute player can be defined as a flute player who plays the flute well. A good will, however, is not to be defined, as Wolff defines it, as a (human) will that wills well. Instead a will (faculty of desire) that wills well is to be defined as a will that takes the idea of a good will (the idea of an autonomous being)²¹ as its object. A ‘good will’ then is an idea, an idea that a human being can (hope to) instantiate if it takes it (the idea) as the object of its faculty of desire. Kant's criticisms of Wolffian rationalism demonstrate that he is not a moral constructivist. An examination of his arguments against Wolff and Baumgarten’s perfectionism draws us to the same conclusion.

(ii) Perfectionism

Kant’s criticism of Wolff and Baumgarten’s perfectionism is similar to his criticism of their rationalism. Once again the problem he points to is that their formulation of the principle of perfection lacks any content. Kant agrees with them that we must aim at perfection, for our moral ideal must be something perfect or holy. Thus, for example, he explains in the *Critique of Practical Reason* that the

law of all laws, like all the moral precepts of the Gospel, presents the moral disposition in its complete perfection, in such a way that as an ideal of holiness it is not attainable by any creature but is yet the archetype which we should strive to approach and resemble in an uninterrupted but endless progress. (5:841)

²¹ Which is the idea of a member of an intelligible world.

Kant, then, does not deny that morality demands that we seek perfection. He believes, however, that the principle “seek perfection” lacks content. Their principle of perfection is merely a negative criterion of morality in the same way that the principle of non-contradiction is merely a negative criterion of truth.

Kant customarily began his lectures on ethics by providing a taxonomy of competing ethical theories, amongst which he includes perfectionism. He traces this position back to Wolff, arguing that, “the principle of perfection, or the harmonizing of the manifold into one, comes from Wolff” (*Ethik Mrongovius*. 29:622). Wolf explains perfection in the following terms:

The agreement of the manifold constitutes perfection. For example, we judge the perfection of a watch by its ability to show correctly the hour and its divisions. It [the watch] is made of many parts assembled together, and these as well as the assembly are aimed at enabling the hands to tell correctly the hour and its divisions. Thus in a watch we find a multitude of things all of which agree with one another. . . The conduct of man consists of many actions and if these all agree with one another, so that they are all finally grounded in one general goal, then the conduct of man will be perfect. (*German Metaphysics*, #157)

Following Leibniz, Wolff argues that perfection involves unity in manifoldness, and he believes that this implies that perfection is relative to kind, for he believes that unity is provided by a thing’s function or goal. What it is to be a perfect watch is different from what it is to be a perfect laptop computer. A perfect watch is one that tells the time perfectly. In such a watch we find unity in manifoldness. The components of which a perfect watch are made are put together in such a way that they all work towards the

same function, telling the time well. Human perfection can be explained in similar terms. Humans have many drives and desires, and do many particular things. A perfect human is one who unifies this manifold of actions and desires, and this can be achieved if they are all directed to one end. Like Kant, then, Wolff maintains that the ethical life involves self-mastery. Kant's objection is that for Wolff it does not seem to matter what goal we set ourselves as long as we follow it rigorously,²² and this seems to run against our ethical intuitions. If, for example, one succeeded in subordinating all of one's inclinations to the goal of being a serial killer one would, presumably, be perfect in Wolff's sense – a perfect serial killer.

Kant makes it clear that this is the reason for his rejection of perfectionism in his lectures on ethics of 1793. Vigilantius, the note-taker, notes that,

In his practical philosophy #39-46, Baumgarten has put forward various formulae which, as imperatives, are supposed to serve for the general principle of all obligation, though professor Kant rejects every one of them. (27:517)

The third of Baumgarten's putative moral principles to be dismissed by Kant is the principle: *Quaere perfectionem, quantum potes* (seek perfection as much as you can), and Kant argues that this supposed principle is,

A formula that contradicts the nature of duty. Perfection is a variable concept. By perfection in general we understand everything we take to be complete, e.g. a perfect liar, a complete villain. . . [etc.]. (27:518)

²² A defender of Wolff, however, could object to such a characterization of Wolff's position, for, like Aristotle, Wolff will argue that human beings have a natural goal (or *ergon*), namely happiness.

The principle of perfection tells us to be good is to be perfect, but a perfect what? This principle is the practical twin of the principle of non-contradiction. Just as the principle of non-contradiction provides us with a merely negative touchstone of truth, the principle of perfection provides us with a merely negative touchstone of morality. We know, Kant believes, that a good will must be perfect, just as a true concept must not contain a contradiction. However, just as the fact that a concept (or proposition) does not contain a contradiction does not mean that the proposition is true (or the concept possible), the fact that something is perfect does not mean that it is good. The principle of morality, the principle we hear when we listen to the voice of conscience, is a principle that provides us with a positive criterion of morally. We recognise, Kant believes, that to be moral is to be perfect, but we also know what type of perfect being we must be.

Kant's own ethics, then, escapes the charge of formalism he levels at Wolff because his aim is to give an account of what the object of 'perfect' volition actually is. Wolff on the other hand fails to explain what the object of a perfectly rational volition would be, and as a consequence his account doesn't end up telling us anything interesting about what is involved in perfectly rational volition. Kant, on the other hand, thinks he is telling us something informative about the nature of perfect rational volition. Perfect volition is volition that has as its goal the idea of being an autonomous agent.

(1e) An Objection to My Interpretation

I have argued that Kant's chief complaint against Wolff and Baumgarten is that their ethics is merely formal. Kant then believes that ethics must have some content. Kant's own position is not formal because he believes that to be virtuous is to make a particular idea the object of one's faculty of desire. What it is to be virtuous can be defined in terms of this idea, and it would not be out of place to call this idea 'the concept of the good', for according to Kant our concept of an autonomous agent (which is the idea of a member of an intelligible world) is our concept of a good will. I have argued that, for Kant, to engage in ethical enquiry is to attempt to clarify theoretically the 'the concept of the good', and that what it is to listen to the categorical imperative is to attempt to instantiate this concept. There are important passages in the *Critique of Practical Reason*, however, in which Kant seems to explicitly argue that the good must be defined with reference to the categorical imperative rather than vice versa, and moral constructivists appeal to these passages to support their interpretation. As I shall show, however, such an interpretation is based on taking these passages out of context, for Kant makes it clear that the priority he gives to the categorical imperative is purely methodological and not ontological or definitional.

Many readers of Kant may be surprised by my characterization of Kant as a moral idealist, for Kant himself is famous for rejecting consequentialism and he himself seems to argue that the foundation of ethics must be formal. Thus, in the *Critique of Practical Reason* he begins a long paragraph by explaining that,

This is the place to explain the paradox of method in a *Critique of Practical Reason*, namely, **that the concept of good and evil must not be determined before the moral law** (for which, as it would seem, this

concept would have to be made the basis) but only (as was done here) after it and by means of it. (5:63 – Kant’s italics, my bolding)

Here Kant explains that we must begin by examining the (purely formal) moral law, and only afterwards define the concept of the good by means of it. Such passages seem to support Rawl’s claim that Kant must be interpreted as a moral constructivist, for they seem to imply that Kant believes the moral law, or CI-procedure, is what makes the concept of the good (say, the idea of being a member of a kingdom of ends) possible. Kant concludes the paragraph by making the same point in even stronger terms, arguing that,

instead of the concept of the good as an object determining and making possible the moral law, it is on the contrary the moral law that first determines and makes possible the concept of the good. (5:64)

Here Kant seems to explicitly deny that the idea of the good can be the basis of ethics. Instead the concept of the good is a secondary notion, parasitic upon the purely formal idea of the moral law, or categorical imperative. However, if we read on, it becomes clear that these passages do not have the import they seem to have at first sight, for Kant himself is quick to qualify these claims, making it clear that the priority he wishes to give to the first formulation of the Categorical imperative is purely methodological. Thus he begins the following paragraph by explaining that the previous “remark . . . concerns **only the method** of ultimate moral **investigation**” (5:64 – my emphasis). These claims, then, about the relationship between the concept of the good and the moral law are merely methodological. Methodologically, we must start by examining the purely formal

notion of duty, and this will point us to the concept of the good, which is the object of duty. This qualification makes it clear that these passages are not meant to imply that the concept of the good is logically or ontologically dependent upon the categorical imperative.

Further support for this interpretation is provided in the *Preface* to the *Critique of Practical Reason*. Here Kant writes that “**among all the ideas of speculative reason** freedom is also the only one the possibility of which we know *a priori*, though without having insight into it, because it is the **condition** of the moral law, which we do know” (5:4 – my emphasis). Kant, then, makes it clear that he believes that the theoretical idea of freedom is a condition of the moral law (or the categorical imperative). And I argue that in these passages when he talks of freedom he means the idea of an autonomous individual.²³ The reason for this is because the categorical imperative orders us to be autonomous, and so the concept of autonomy is a condition for the possibility of the imperative. In a famous footnote to this passage, Kant explains the relationship between the concept of freedom and the moral law in more detail. He writes:

Freedom [and I am assuming that Kant here means ‘the idea of an autonomous individual’] is indeed the *ratio essendi* of the moral law [the categorical imperative]; the moral law is the *ratio cognoscendi* of freedom. (5:4)

Our consciousness of the categorical imperative, then, is the *ratio cognoscendi* of the concept of the good, but the concept of the good is the *ratio essendi* of the categorical imperative. What it is to be virtuous is to be defined in terms of the object of volition. For

²³ And not merely the idea of freedom in the negative sense.

Kant, to be virtuous is to strive to be an autonomous being; this is what the categorical imperative demands. The idea of an autonomous being is an ideal and, as such, it is the idea of a holy being. Our awareness of the categorical imperative involves a consciousness of this ideal, albeit an obscure one. Autonomy cannot be defined in terms of acting out of duty, for, as we have seen, Kant believes that there is nothing contradictory in the idea of an autonomous being that does not act upon imperatives. Rather, it is our duty to be an autonomous agent. This is what Kant means when he claims that, “freedom is the *ratio essendi* of the moral law”. Because the categorical imperative demands that we be autonomous, to avoid circularity in our definition, the idea of an autonomous being (a good will) must be defined without reference to the notion of duty, and Kant attempts to provide such a definition in *Groundwork* II. The idea of a good will is the idea of a member of a kingdom of ends (or intelligible world) and this idea is an idea of pure reason. This idea is, as we have seen, the idea of what Kant elsewhere calls a holy being, and so we can say that although virtue is the *ratio cognoscendi* of holiness, holiness is the *ratio essendi* of virtue. For, what it is to be virtuous is to strive for holiness. This is why holiness is the *ratio essendi* of virtue, because if we did not have the idea of holiness we could not strive towards it. Virtue is the *ratio cognoscendi* of holiness because the voice of conscience makes us immediately aware of our duty, but provides us with an obscure representation of the object of our duty, which is to be holy. The imperative commands us to be holy. As such it must involve a consciousness (albeit and obscure on) of what it is to be holy. An investigation, then, of the demands the categorical imperative presents us with can help us clarify what is involved in our idea of a holy being.

In claiming that Kant is a moral constructivist, Rawls is, in effect, claiming that Kant believes that the CI-procedure is both the *ratio cognoscendi* and the *ratio essendi* of the idea of a kingdom of ends. He confuses what is a methodological issue with an ontological or logical issue. Kant begins the *Groundwork* examining the procedures a virtuous will (implicitly) acts upon. He begins by examining what Rawls calls practical rationality, and he reaches the conclusion that a person who is practically rational is striving to be a member of a kingdom of ends. This, however, is merely Kant's method of investigation, and he does not draw the conclusion that this fact about the appropriate order of investigation implies that the idea of (a member of) a kingdom of ends is logically dependent upon an understanding of practical rationality (the CI Procedure), as Rawls believes. Instead, in claiming that freedom is the *ratio essendi* of the moral law, Kant means that practical rationality is to be defined in terms of autonomy. A practically rational human being is defined as a human being who strives to be an autonomous individual, and to avoid circularity we must be able to give an account and definition of what it is to be an autonomous individual without appealing to the notion of practical rationality. We can do this, I believe, by examining the concept of community, and this was one of the principle goals of my dissertation.

(1f) The Value of Moral Philosophy

The moral philosopher is concerned with providing theoretical clarification of our moral ideas. The question I wish to address in the remainder of this paper is: what is the

practical value of such an undertaking? The reason this is a pressing question is because Kant believes that we do not need a clear theoretical understanding of our moral concepts in order to be virtuous. The obscure consciousness provided by our conscience is perfectly adequate as a guide to action; if we listen to our conscience we cannot fail to be virtuous. What then is the value of a theoretical investigation of our moral concepts? I will argue that although we do not need conceptual clarity in order to be virtuous, such clarification can help us to be virtuous in so far as it can help us to really listen to the voice of conscience.

Many readers of Kant's ethical writings take him to be primarily concerned with offering guidelines for action. At the very least, they write about Kant as if this were the purpose of his ethical writings. For example, Korsgaard (1996), in her influential article on Kant's strategy in *Groundwork I* writes that

The argument of *Groundwork I* is an attempt to give what I call a "motivational analysis" of the concept of a right action, in order to **discover** what that concept applies to, that is, **which actions are right**. (p.47 – my emphasis)²⁴

This, however, is a fundamentally misguided way of reading Kant, for he repeatedly asserts that we do not need to do moral philosophy in order to "discover" which actions are right. We all already know how to behave morally, and so do not need philosophers to tell us this. "Common human reason", Kant argues in the *Groundwork*, "knows very well how to distinguish in every case that comes up what is good and what is evil, what is in

²⁴ I'm not sure that Korsgaard actually meant to say this, and I suspect that upon reflection she would have chosen her words more carefully, however she does say it, and similar remarks abound in the secondary literature, and even if they do not necessarily reflect the authors considered opinion, they do suggest that the author has a certain conception of the function of moral philosophy at the back of their mind.

conformity to duty or contrary to duty” (4:404). The pre-philosophical understanding, then, knows perfectly well how to act morally. Unlike the Wolffians, and many other traditional moral philosophers in the western tradition, Kant is not an intellectualist; he denies that immoral behavior is the result of mistaken beliefs about the good. It is instructive to compare Kant’s conception of moral philosophy with Wolff’s. Wolff is an intellectualist and believes that if we knew what was good for us we would do it. He believes that there is no gap between recognizing an act as good and willing it. Thus, he can claim that:

The knowledge of good is a motive [*Bewegungsgrund*] of the will. . . it cannot happen that one does not will an inherently good act if one distinctly conceives it. . . So if we do not will [such acts], there is no other cause than that we do not recognize them [as good]. (*Reasonable Thoughts about the Actions of Men*, #6)

As a result of this Wolff maintains that the only way we can become better human beings is by improving our knowledge of the good. Unfortunately, discovering the truth about the good is an arduous task, and Wolff is worried that “perhaps someone will wonder how it will go with the pursuit of good and the omission of bad if so much is required in order to distinguish good from bad” (*Reasonable thoughts about the Actions of Men*, #150). His response is to suggest a division of labor. It is not necessary that all people spend time putting in the work to distinguish the good from the bad; this job can be left to the philosophers. Thus Wolff writes that,

Here it will do to respond that we are not speaking only of those who are to generate from their own reflections the rules according to which men are to judge their free actions in different conditions of life, that is, of the discoverers of the truths that belong to a doctrine of morals. But it is not necessary that all men be discoverers. It is enough if some among the

learned devote themselves to discovery, whose findings the others can afterwards learn, which is much easier. (*ibid.*)

According to Wolff, then, the function of the moral philosopher is to discover the truth about good and bad and to communicate this truth to those who do not have the time, inclination or capacity to think about such matters.²⁵ The function of moral philosophy is, from this perspective, to communicate moral knowledge to the ignorant so that they can lead more virtuous lives.

Kant, in contrast, has far more respect for the moral capacities of common men and women, and as a result of this he believes that there is “no need of science and philosophy to know what one has to do to be honest and good, and even wise and virtuous” (*Groundwork*, 4:404).²⁶ If, however, we do not need to do moral philosophy to know what one has to do in order to be good, why do we need moral philosophy? Any serious interpretation of Kant’s ethics must be able to answer this question, for Kant believes that the moral philosopher does have an important role to play.

Intellectualism and Voluntarism

²⁵ Wolff argues that this is the primary reason he writes his ethical works. Thus, he writes: “Because not everyone is skilled in discovery and because one is not to live for oneself alone but to care also for others, these to whom God has lent the strength and opportunity to carry out this work better than others are obligated to share in books for the others what they have learned about the soul. From this urge there came what I have written in my *Reasonable Thoughts about God the World and the Human Soul* and what else I shall put in this book” (*Reasonable Thoughts about the Actions of Men*, #233).

²⁶ This respect for the moral capacities of common people may be traceable to Rousseau’s influence, thus Kant writes in 1763 that, “there was a time. . . when I despised the masses, which knew nothing. Rousseau has set me right. This blind prejudice disappears; I learn to honor men. . .” (8:624). Schilpp (1938), however, in his influential study on Kant’s pre-critical ethics, suggests that Kant had probably already been “indoctrinated with the notion of the inherent worth of every human being as a youthful Pietist” (p.49).

To understand Kant's account of the role and value of moral philosophy it is necessary to understand his account of moral motivation and his explanation of the possibility of immorality. He wishes to steer a middle course between the Scylla of voluntarism and the Charybdis of intellectualism. The intellectualist maintains that the good is irresistible, for we always act under the aspect of the good. As a result of this the intellectualist claims that all immoral behavior is a result of not having a clear enough understanding of the good. The voluntarist, in contrast, maintains that, however clearly we understand the good we are always free to choose the bad. Kant agrees with intellectualist that the good is irresistible. However, he also agrees with the voluntarist that we are free to choose the bad. And, in addition, he believes that all of us, possessing a conscience, have an adequate understanding of the good. On the face of it these commitments seem to be incompatible, for if our conscience provides us with an adequate understanding of what is involved in being good, and the good is irresistible, it is not clear how, and in what sense, we can choose to be bad. Kant's solution will be to argue, against the intellectualist, that our consciousness of the good is not directly through the intellect, but through the conscience.²⁷ The judgment of conscience is, Kant maintains, infallible and, *in so far as we choose to listen to it*, irresistible. As a result of this, if we were truly attentive to the voice of conscience we would always behave morally. We often, however, fail to pay attention to this voice, and this failure is, Kant believes, always the result of a free choice. In so far as we pay attention to the judgment of conscience we will be virtuous; we are

²⁷ Although our consciousness of the 'good' is through the conscience and not the intellect, I will argue that for Kant, the object of the conscience is intelligible, being an idea of pure reason that presents itself to our conscience as a practical *ideal*.

free, however, not to listen. This is the locus of our freedom.²⁸ The moral law, then, is like the Siren's voice; in so far as we choose to listen to it, it is irresistible. We are free, however, to plug our ears or to drown it out with competing voices.

This account of the locus of free choice helps us understand Kant's conception of the role of moral philosophy. As we have seen, Kant claims that his goal in the *Groundwork* is to "clarify" the concept of a good will. If Kant were an intellectualist, realizing this goal would serve an obvious moral function. For the intellectualist believes that immoral behavior is always, and only, a failure to truly understand the good, and so the clearer our understanding of what is good, the better we will be. Kant, however, is not an intellectualist; indeed, he believes that even the pre-philosophical human understanding has a good enough understanding of the good to always act rightly. For Kant, knowing the good and doing the good are two distinct things. Kant himself claims that he wishes to clarify our understanding of the concept of a good will so that it will "shine forth more brightly". And this, I believe, provides us with a clue for understanding why Kant believes that engaging in moral philosophy has some practical value. For the brighter something shines, the more noticeable it is, the harder it is to fail to pay attention to it. Engaging in moral philosophy, and clarifying the idea of a good will, helps us to pay attention to the moral ideal. It does this in two ways. Firstly, in making the concept of a good will "shine forth more brightly", it amplifies the volume of the judgment of conscience. Secondly, and perhaps more importantly, it makes it harder for us to give ourselves excuses; for excuses are, Kant believes, like voices that compete with the voice

²⁸ And here I am talking about what Kant calls freedom in the negative sense.

of conscience for the attention of our will. This explains the practical value of engaging in moral philosophy.

The court of conscience

The primary role of the moral philosopher, then, is to provide us with a tool that we can use to help ourselves pay attention to the judgment of conscience. For Kant the conscience is like an internal court that operates both prior to and after our actions. Thus he explains in his ethics lectures that, “consciousness of an internal court in a human being (“before which his thoughts accuse or excuse one another”) is **conscience**” (27:168), and he distinguishes between what he calls a *conscientia antecedens* (which operates before we act) and a *conscientia consequens* (which comes into operation after we act). This conscience is like a court and involves,

The assumption of an accuser, who seeks to arouse the conscience; a defender, who tries as an advocate to assuage it; and a judge, who assesses the action by the laws of duty and establishes the consequences. (*ibid.*)

When we behave, or plan to behave, immorally we fail to listen to the judgment of this court. This judgment, Kant believes, cannot err: the notion of “an erring conscience”, he writes in the *Metaphysics of Morals*, “is an absurdity” (6:401).²⁹ When we behave immorally it is not that our conscience has erred. Nor is it the case that our conscience has failed to make a judgment, for the judgment of conscience is “an unavoidable fact” (6:400). “Unconscientiousness”, then, “is not lack of conscience but propensity to pay no

²⁹ And “if someone is aware that he has acted in accordance with his conscience, then as far as guilt or innocence is concerned nothing more can be required of him” (6:401). See also, *Ethik Collins*, 27:354-5.

heed to its judgment” (6:401).³⁰ When we behave immorally we attempt to give excuses to ourselves before this court for not living up to our moral ideal; we choose to listen to the defense attorney and try to block out the voice of the judge.³¹

The voice of the judge of conscience is like the voice of a siren; in so far as we listen to it, it is irresistible. Kant calls this irresistibility of the voice of the judgment of conscience our “moral predisposition”. As a result of our predisposition to morality, Kant believes that we can only fail to be guided by the voice of our conscience if we stop up our ears or drown it out with other voices. Thus, in the *Religion*, Kant argues that,

The human being (even the worst) does not repudiate the moral law, whatever his maxims, in rebellious attitude (be revoking obedience to it). **The law rather imposes itself on him irresistibly, because of his moral predisposition;** and if no other incentive were at work against it, he would also incorporate it into his supreme maxim as sufficient determination of his power of choice, i.e. he would be morally good. (6:36 – my emphasis)

The judgment of conscience is irresistible and, insofar as our inner ear is open, our faculty of desire will be determined [*bestimmt*] by the law.³² Morality, for Kant, like the Stoic, is a matter of attentiveness, and immorality is the result of a, freely chosen, distraction. Engaging in moral philosophy helps us avoid distraction by “providing access” (4:437) to the law.

³⁰ “So when it is said that a certain human being has no conscience, what is meant is that he pays no heed to its verdict” (6:400).

³¹ As Kant points out, “The defending or consoling conscience may work very much to our disadvantage” (27:619).

³² It is interesting to note the prevalence of aural metaphors in Kant’s ethics, and in general in his account of reason. The primary activity of reason is, I believe, what Kant calls “determination” [*Bestimmung*], and the etymologically it is related to “*Stimme*” (Voice). In the eighteenth century, “*Bestimmung*” could also mean “vocation” or “calling”.

Having said this, the purely intellectual activity of clarifying the idea of a good will does have some practical effect, for the clearer our theoretical understanding of the moral ideal, the more difficult it becomes for us to give ourselves excuses for not living up to it in practice. The reason for this is that the excuses we provide ourselves with in the court of conscience are themselves products of the intellect. Clarifying the concept of a good will can help us quiet the voice of the defense attorney in this court, and it is this voice that distracts our attention from the voice of the moral judge.

Kant believes, then, that although the moral ideal infallibly serves as a “compass” for our actions, and one that we all have “always before our eyes” (4:403), we often deliberately squint when we are taking directions. The moral philosopher cannot stop us squinting; he can, however, point out to us that we are squinting, which makes it more difficult for us to give excuses to ourselves if we choose to head in the wrong direction. The purpose of the moral philosopher, then, is to help us focus our attention on the moral ideal. Clarifying the moral ideal is to amplify the voice of the judge within, which is in danger of being drowned out by the pleading voice of the defense attorney in the court of conscience. Engaging in moral philosophy, then, can help us promote a particular duty, namely,

the duty to cultivate one’s conscience, to sharpen one’s attentiveness to the voice of the inner judge and to use every means to obtain a hearing for it.
(*Metaphysics of Morals*, 6:401)

An Analogy

Kant believes that we all know what it is to do the right thing. Even the least educated peasant knows *how* to be moral; the judgment of conscience is the same for him as for the enlightened philosopher. The problem with the philosophically uneducated is that they do not clearly grasp theoretically *what* it is to be moral. Their representation of *what* it is to be moral is 'obscure'. This obscure representation of what it is to be moral is a good enough guide to action. This obscure representation is what provides the common human understanding with its moral *know-how*.

Imagine going for a walk in the country with a friend. At one point you decide to split ways and your friend points to a building in the distance, a house on a hill, and tells you that he will meet you there in three hours. You vaguely make out something in the distance through the mist, but it is too far away to make out what it is. But you know that if you pay attention to where you are going you will have no problem finding your way there. You wander off and become distracted by the view and the fresh country air, and stop paying attention to where you are heading and suddenly there are two hills in front of you, one with a house on top and one with a pub. And you end up waiting for your friend in the pub, quite happily drinking a few beers with the locals, occasionally checking your watch and wondering where he is. Just before closing time he storms into the pub, soaking wet, having spent the last three hours waiting for you, outside the house. And he's not happy. You tell him that you're really sorry, but that he shouldn't blame you or be angry with you because you honestly thought you were in the right place, and in some sense it is true that you didn't *know* that you were waiting in the wrong place. Trying to convince your friend, or yourself, that you are entirely blameless, however, is

in bad faith. It is true that you didn't choose to make it your maxim to stand up your friend.³³ If someone had come into the pub and told you that your friend was waiting outside in the rain for you, you would have jumped up and gone to find him. Your being in the wrong place, however, is willful in that you chose not to pay attention to where you were going and your friend can legitimately be angry with you, for having (at least in some sense) violated some ideal of friendship. If you had really cared about (meeting) him you would have paid attention to where you were going.

What Kant calls the common human understanding has an analogous relationship to the idea of a good will as my traveler has to the house on the hill. The common human understanding knows how to be moral just as my traveler knows how to get to his meeting point. The common human understanding, however, cannot really give an adequate description of what it is to be moral. And this makes it a lot easier for it to give itself excuses when it fails to reach its destination. Kant's purpose in the *Groundwork* is to exhibit this pure idea (the idea of a good will) in all its purity. He wants us to be in a position that when we part ways, and arrange to meet, our friend can tell us: "let's meet outside that house over there, on the hill opposite the King's Arms pub". I can still end up in the pub instead of waiting outside the house, but if I've been given such instructions it's much harder for me to give my friend, or myself, excuses. If my friend has given me such clear instructions it will be much harder to convince him, or myself, that I honestly cared about the meeting if I end up in the wrong place. This analogy helps us understand

³³ And my inattentive traveler is analogous to Kant's conception of an (immoral) human being in this regard too. For Kant does not believe in the possibility of a diabolic will. That is, he does not believe in the possibility of a will governed by an evil principle. See the discussion of the notion of a diabolic will in *Religion*.

the role of the moral philosopher and teacher. The role of the moral teacher is not to tell us *how* to be moral; we already know that. The role of the moral educator is to make our obscure understanding of *what* it is to be moral shine out more brightly. This clearer understanding will not make us moral, but it will make it harder for us to listen to the sweet words of the defense attorney in the court of conscience.

Kant often suggests that there are two principal excuses we give ourselves when we act, or plan to act, immorally. Engaging in moral philosophy, makes it harder for us to give ourselves these excuses, for these excuses always have a basis in the intellect and thus can be combated by theoretical enquiry.

(1) The first type of excuse is a result of a willful **corruption** of our pure moral ideal.³⁴
(2) The second type of excuse involves willfully convincing ourselves that our moral ideal is impossible, either in itself or as an idea that can serve as an ideal for beings like us. This is **despair**.

(1) In the first case when we choose to act immorally we, in effect, choose to be guided not by a pure ideal but by a watered down version of it. This is what happens, Kant believes, when we take happiness as our ideal³⁵, or when we take another particular individual as our ideal, comparing ourselves not to the pure idea of a good will, but with

³⁴ Thus Kant writes that “a metaphysics of morals is. . . indispensably necessary, not merely because of a motive to speculation – for investigating the source of the practical basic principles that lie a priori in our reason – but also because morals themselves remain subject to all sorts of corruption as long as we are without that clue and supreme norm by which to appraise them correctly.” (*Groundwork*, 4:389-9).

³⁵ For happiness is an ideal, but, as Kant explains in the *Groundwork*, it “is not an ideal of reason but the imagination” (4:418). An account of the way in which happiness is an imperfect ideal is beyond the scope of this dissertation.

another, imperfect, human being, and taking that person's imperfect behaviour as our standard. Choosing to imitate another person rather than taking the pure idea of an autonomous individual as our moral archetype, is one of the most common causes of excuses.³⁶ For it allows us to excuse all sorts of behavior as 'excusably' human. Conceptual clarity about what is involved in our *pure* idea of a good will makes it easier for us to avoid the temptation of substituting the corrupted pseudo-ideal of happiness for the pure idea of a good will or of imitating others.

(2) In the second case we willfully convince ourselves that it is impossible for us to live up to the ideal, either because we convince ourselves that the idea of an autonomous agent is itself impossible, or because we convince ourselves that it is impossible for creatures such as us to live up to such an ideal. There are two forms of despair. The first type of despair is a result of a recognition of our weakness as sensuous beings who have needs. When we give in to such despair we tell ourselves that there are certain needs we have that it is *physically impossible* not to satisfy.³⁷ The second type of despair is a result of the fact that we necessarily experience ourselves as phenomenal spatio-temporal (and hence conditioned) beings, whereas our moral ideal is an idea of pure reason, and as such

³⁶ Thus Kant explains, in his ethics lectures he gave at the time he was working on the Groundwork, that, "an example is when a general proposition of reason is exhibited *in concreto* in the given case. . . No examples are needed in matters of religion and morality. There is thus no pattern in religion, since the ground, the *principum* of behavior, must lie in reason and cannot be derived *a posteriori*. . . The examples, therefore, must be judged by moral rules, not morality and religion by the examples. The archetype lies in the understanding" (*Ethik Collins*, 27:332-3). One of the targets of this attack is almost certainly the pietist culture around him. Perhaps the most popular genre of 18th century pietist literature was the conversion narrative. In these autobiographical narratives pietists told the story of their own rebirth. Such narratives were extremely popular and were presented as models to be followed. Thus Semler (1781), a contemporary of Kant's, could write that, for the pietists "the story of ones own experience and edification became the rule to follow exactly" (quoted from Fulbrook 1983, p.171).

³⁷ Often such despair is the result of inauthenticity in the heideggerian sense. Instead of saying: I have chosen to do x, we tell ourselves that 'one' acts like this in these circumstances.

it is the idea of a being that is unconditioned. When we give in to this form of despair we tell ourselves that it is *logically impossible* for a phenomenal being to be an autonomous individual. There is a natural tendency in human beings towards such despair, for there is a natural tendency to believe that every object of experience (including ourselves) must be causally conditioned and divisible. For every object experienced in space/time is experienced as infinitely divisible and causally determined. Our moral ideal, however, is the idea of an autonomous individual, and the idea of an autonomous individual is, as autonomous, the idea of a being that is not causally determined, and, as an individual, it is the idea of a being that is simple, in the sense of being indivisible. Hence no object of experience can be experienced as an autonomous individual. Indeed, if we believed that space and time were conditions for the possibility of objects, and not merely conditions for our experience of objects, we would be justified in concluding that no object (of experience) could possibly be an autonomous individual. If this were the case, then our despair would be justified. Kant, however, believes that transcendental idealism offer a way out of this despair, for the transcendental idealist maintains that just because we must experience objects as essentially causally conditioned and divisible this does not imply that we cannot *think* of objects of experience (and, in particular, of ourselves and other human bodies) as autonomous individuals without contradiction. Another function of the moral philosopher, then, is to explain how it is possible for us to think of phenomenal objects (including ourselves) as autonomous individuals without contradiction.

The moral philosopher can help us avoid both forms of despair. He can help us avoid the first kind of despair by convincing us that we are free and that, in addition, it is not unreasonable to hope that there is a just and benevolent God who ensures that happiness is allocated in proportion to virtue, so if we rigidly follow our duty it is not unreasonable to hope that our needs will be met. He can help us avoid the second type of despair by convincing us that although we *intuit* ourselves as essentially causally determined (conditioned) beings, there is nothing contradictory in *thinking* of ourselves as autonomous (unconditioned) individuals. The moral philosopher can do this by convincing us of the plausibility of transcendental (or formal) idealism. Formal idealism makes a distinction between the form and matter of the objects of intuition, and claims that although the matter of the objects of intuition is real, the (spatio-temporal) form of these objects is ideal and subjective. If we accept transcendental idealism it allows us to *think* of objects of intuition, including ourselves, as unconditioned without contradiction. For although an object of experience cannot be thought of as unconditioned and simple *qua* object of experience there is nothing contradictory in thinking of such an object as unconditioned and simple in itself.³⁸

To conclude: Engaging in metaphysical speculation and working out clearly what is involved in the *idea* of a member of an intelligible world, and recognizing that there are no contradictions involved in it, cannot make us virtuous; to be virtuous, all we need to do is to listen to our conscience. Kant believes that grasping the idea of a good will in its purity is neither necessary nor sufficient for us actually to be good. Having grasped the

³⁸ This argument will be examined in more detail in chapter five.

idea in its purity we are still faced with the choice of whether or not to recognize it as a practical ideal and to choose to instantiate it. Grasping the idea in its purity, however, can make the choice to be good easier, or at the very least, it makes it more difficult for us to give ourselves excuses for not living up to it. Thus, the conceptual clarity that arises as a result of engaging with traditional metaphysical questions about the nature of the intelligible world cannot provide us with knowledge, but it can help us avoid *temptation* and *despair*. It can help us avoid the temptation of replacing our *pure* idea with a corrupted watered-down impure, pseudo-ideal (such as the ideal of happiness) and it can help us avoid the despair of convincing ourselves that there is no such ideal or that it is impossible for us to live up to or instantiate such an ideal.

To summarize: (1) Kant's ethics is grounded on the pure idea of a pure will, and this is the idea of a holy will. And, (2) this idea is pure in the sense that it is an idea of the pure intellect and as such it can be completely **determined** (in the sense of being defined) purely in terms of *a priori* principles and concepts. (3) We all already possess this idea and our consciousness of duty involves an obscure and subjective awareness of this idea as the standard of our action. (4) This obscure awareness of the idea of a good will is analogous to a 'savage's' awareness of a house. The 'savage' sees the house well enough to be able to find his way there, but he does not know what it is. Similarly the 'common human understanding' or innocent pre-philosophical individual 'sees' the idea of a good will when she is aware of her duty, and this pre-philosophical grasp of the idea is good enough as a guide for action. Such an understanding, however, does not have a clear understanding of this idea. (5) The function of the moral philosopher is to help us

understand clearly what the idea of a good will is an idea of. Thus, Kant's primary aim in his ethical writings is to clarify the idea of a good will, in the sense of making us conscious of what pure (theoretical) idea the idea is. In so doing Kant hopes to lead us from a subjective consciousness of this idea to an objective one, from an obscure consciousness to clear understanding. (6) The reason this is important is that the clearer our understanding of the moral ideal (which is the object of duty), the harder it is for us to give excuses to ourselves for not living up to it.

In his lectures, Kant explains that,

Necessitation is conceivable only where a contravention of moral laws is possible, and hence a thing can be morally necessary without being a duty, which would happen if the subject were at all times to act without necessitation in accordance to the moral law; for then a duty or obligation so to act would not be present; hence this does not hold of **a morally perfect being**, in that such a being acts solely from holiness. . . . Where there is no necessitation, there also no moral imperative, no obligation, duty, virtue, ought or constraint is conceivable. Hence the moral laws are also called laws of duty, because they presuppose an agent subject to the impulses of nature. . . . **Like an angel, a being of this kind** [a morally perfect being] **can in no way be thought of existing, but to the philosopher is merely an idea.** (*Ethik Vigilantius*, 27:489-9 – my emphasis)

My claim, then, is that Kant believes that for us to be moral (virtuous), is to choose to be such a morally perfect being. It is, in other words, our duty to become such a being. A morally perfect being cannot be thought of as existing, at least as an object of experience, but we do have an idea of such a being. The bulk of this dissertation will be devoted to justifying this claim, and explaining in detail what this idea is an idea of. I argue that this idea is an idea of pure reason, and in particular an idea of rational cosmology, for it is the

idea of a member of an intelligible world. And **the idea of an intelligible world can only be the idea of a political community of autonomous individuals subject to juridical laws.** Our idea of a holy will is the idea of a member of such a community. This may, at this stage, seem a pretty odd claim. Hopefully, by the end of this dissertation it will make more sense.

In addition to explaining the moral ethical position of the mature Kant, I will also give an account of how his position developed. In the previous citation, Kant identifies the idea of a morally perfect being with the idea of an angel. Thus, if my interpretation is correct, one could say that the highest moral imperative is: be an angel! This seems a pretty bizarre claim to make. It was, however, precisely the moral position of the eighteenth century Swedish mystic and visionary Emanuel Swedenborg. Swedenborg believed he had had his spiritual eyes opened by God and that he had been allowed to visit both heaven and hell. He describes heaven as a community of angels governed by moral laws (of benevolence) and believes that all angels were at one time men, and that we are ourselves members of a spiritual community, and it is our choice whether to join the community of heaven or the community of hell. This choice, Swedenborg believes, is the choice of our moral character, and a choice that we must make in this world. Kant read Swedenborg in the early 1760s, and this was time most commentators believe he first began to formulate his mature ethical position, and in 1766 he published a book on Swedenborg, *Dreams of a Spirit Seer*. I will argue that Kant was deeply influenced by his engagement with Swedenborg, for in the process of his engagement with the Swedish seer he to came to conceive of morality in terms of being a member of a (spiritual)

community. I suggest, then, that Kant's idea of a kingdom of ends has its genesis in Swedenborg's vision of heaven as a community of spirits governed by divine laws of benevolence. In chapter three I will attempt to justify this claim. My account of Kant's development, if correct, suggests that, at least historically, the idea of being a member of a kingdom of ends is more basic to Kant's ethics than the idea of being autonomous. Indeed, I will argue that autonomy is ultimately important for Kant because the only way we can conceive of a kingdom of ends, as a real community, is as a community of autonomous agents. Autonomy, then, is important for Kant in the sense of being part of the citizenship requirements for a possible kingdom of ends.

Although the genesis of Kant's position can be traced back to his engagement with Swedenborg, Kant's ultimate ethical position is significantly different. Firstly, although Swedenborg conceives of heaven as a community of spirits governed by moral laws, he does not suggest these laws must be given by the members of the community themselves. In Chapter four I will explain why Kant came to see that our idea of such a community must be the idea of a community of autonomous individuals. The heart of his argument is that a community can only be really unified, and hence can only be *a* community, if the individual members of the community are the source of the laws that provide the community with its unity. In other words a community can only be *a* community if the members of the community are autonomous.

Secondly, Swedenborg conceives of heaven as a community of angels governed by laws of love, or what Kant would call laws of benevolence. The mature Kant, in contrast,

conceives of the kingdom of ends as a political community, governed not by laws of benevolence, but by juridical laws. Thus, in the *Religion* he praise Jesus as a moral teacher and explains that,

The teacher of the Gospel manifested the Kingdom of God on earth to his disciples only from its glorious, edifying and moral side, namely in terms of the merit of being citizens of a divine state; and he instructed them as to what they had to do, not only that they attain to it themselves, but that they be united in it with others of like mind, and if possible with the whole human race. (6:135)

Our idea of a good will, or holy individual, is, I suggest, the idea of a “citizen of a divine state”. To be virtuous is to strive to be a member of such an ideal state. The categorical imperative tells us what we have to do to be eligible for citizenship in such a state.

In Chapters seven, eight and nine, I will defend the claim that, for Kant, the idea of a Kingdom of ends is the idea of an ideal state. I will explain how he believes it is possible for us to have the idea of such a state *a priori*³⁹, and why we can only conceive of such a state as governed by juridical laws.

³⁹ I will argue that that this idea can be derived from the category of community, which is based upon the disjunctive form of judgment.

Chapter Two

Moral Idealism & Kant's Ethic of Interaction

Central to my interpretation of Kant's ethics is the belief that Kant should be understood both as (a) a 'moral idealist' and as (b) advocating an 'ethics of interaction'. I call Kant's ethics an 'ethics of interaction' because Kant's moral ideal is the idea of being an individual member of a world of individuals, and, for Kant, a world is only really a world, as opposed to a mere aggregation, if the individual beings that constitute the world are in interaction. The focus of this chapter will be on explaining and defending my claim that Kant's moral ideal is the idea of being a member of a/the intelligible world.

(2a) Kant's Moral Idealism.

The ideal man

By labeling Kant a moral idealist I mean that a pure ideal stands at the heart of his ethics and that to be moral is (or would be) to take this ideal as the object of our choice; it is our duty to become such an individual. Kant makes this clear in his lectures on ethics from

1785, a lecture course given around the time of the publication of the *Groundwork*. Here he explains that,

to expound morality in its full purity is to set forth an Idea of practical reason. Such Ideas are not chimeras, for they constitute the guideline to which we must constantly approach. . . We have to possess a yardstick by which to estimate our moral worth, and know the degree to which we are faulty and deficient. . . An ideal is the representation of a single thing, in which we depict such an idea to ourself in *concreto*. All ideals are fictions. We attempt, in *concreto*, to envisage a being that is congruent with the idea. In the ideal we turn the ideas into a model. . . The ideal is a *prototypon* of morality. (*Ethik Mrongovius*, 29:604-5)

Elsewhere in the same lectures he argues that, “The principle of morality is. . . the Idea of a will, insofar as it is a law unto itself” (29:628). This practical archetype, then, is the idea of an autonomous agent. In other passages he argues that the moral ideal, the yardstick by which we estimate our moral worth, is the idea of an ideal man. And he explains that,

We conceive of man first of all as an ideal, as he ought to be and can be, merely according to reason, and call this Idea *homo noumenon*; this being is thought of in relation to another, as though the later were restrained by him; this is man in the state of sensibility, who is called *homo phenomenon*. The latter is the person, and the former merely a personified idea; there, man is simply under the moral law, but here he is a phenomenon, affected by the feelings of pleasure and pain, and must be coerced by the noumenon into the performance of duty. (*Ethik Vigilantius*, 27:593)⁴⁰

⁴⁰ Kant makes a similar point in his lectures from 1793. Following standard natural law theory he argues that obligation requires both an obliging (*obligans*) and obliged (*obligatus*). Thus he writes that, “Although. . . obligation is established by reason, it is nevertheless assumed that in the performance of our duty we have to regard ourselves as passive beings, and that another person must be present, who necessitates us to duty”. Typically the source of the law was found in God. “Crusius found this necessitating person in God, and Baumgarten likewise in the divine will, albeit known through reason, and not positively. . . .” Kant, however, argues that we do not need to appeal to God in order to explain obligation, for “if **the obligator is personified as an ideal being or moral person**, it can be none other than the legislation of reason; this, then, is **man considered solely as an intelligible being**, who here obligates man as a sensory being, and we thus have a relationship of man qua phenomenon towards himself qua noumenon” (*Ethik Vigilantius*, 27:510 – my emphasis). Similarly in 1784-5 he argues that, “We have in ourselves two foundations for our actions: inclinations, which are animal in nature, and humanity, to which the inclinations have to be subordinated” (*Ethik Collins*, 27:347). Here he suggests that our animal inclinations have to be

Kant, then, would endorse Schiller's claim in his *Letters on the Aesthetic Education of Man*, that,

Every individual human being, one may say, carries within him, potentially and prescriptively, an ideal man, the archetype of a human being, and it is his life's task to be, through all his changing manifestations, in harmony with the unchanging unity of this ideal. (Schiller 1967, p.17)

To be moral, then, is to attempt to live up to this ideal. But what is this ideal? This question will be the main topic of this dissertation. I will argue that Kant believes that our idea of an ideal man is the pure idea of an individual member of an intelligible world. And Kant believes that our idea of an intelligible world is the pure idea of a community of individuals in interaction. For Kant, then, to be moral is to choose to be a member of a community and to interact really with others.

Our idea of a community is a pure theoretical idea derived from the category of community, the third category of relation in Kant's table of categories. I will argue that Kant believes that the only way we can conceive of a community is if we think of the individuals that constitute the community as autonomous agents. As a result of this the idea of an individual member of a community *is* the idea of an autonomous agent. So to choose to be a member of a community is to choose to be autonomous.

The distinction between *a priority* and purity

subordinated to our (idea of) humanity. Note that the relationship of reason to inclination is one of subordination not eradication.

Kant argues that the moral ideal must be a *pure* idea. Kant stresses the importance of the *purity* of ethics in all of his mature moral writings, so it is important to understand what he means by this term. In his logic lectures, he explains that,

A concept is either an *empirical* or a *pure* concept. A pure concept is one that is not abstracted from experience (*vel empiricus vel intellectualis*). A *pure* concept is one that is not abstracted from experience but arises rather from the understanding even *as to content*. (*Jäsche Logic*, 9:92)

At first sight it might seem that in contrasting pure with empirical, Kant simply wishes to identify purity with *a priori*. Purity, however, involves more than mere *a priori*, for a pure concept must not only be *a priori*, but it must also “arise from the understanding as to its content”. Kant is more explicit about the distinction between purity and *a priori* in his introduction to the *Critique of Pure Reason*. In the first edition introduction he offers a definition of purity, explaining that, “every cognition is called **pure**. . .that is not mixed with anything foreign to it” (A10/B24).⁴¹ In the second edition he expands on this and makes it clear that he believes that not all *a priori* cognition is pure. Thus he explains that,

Among *a priori* cognitions, however, those are called **pure** with which nothing empirical is intermixed. Thus, e.g., the proposition “Every alteration has its cause” is an *a priori* proposition, only not pure, since alteration is a concept that can be drawn only from experience. (B3)

The example Kant gives here is of a *proposition* that is *a priori* but not pure, but his account of why it is not pure can also be applied to *concepts*. Indeed, the reason why the proposition Kant appeals to is not pure is because the concept of alteration is itself not a

⁴¹ Kant’s emphasis.

pure concept. The unschematised concept of a cause is one of the categories, and so is a product of the pure intellect; it is *a priori* and pure.⁴² The notion of alteration, on the other hand, although it is *a priori*, is not a product of the pure intellect, for the notion of alteration only makes sense in time, and so involves both the intellect and the faculty of intuition. Thus, Kant explains in the *Transcendental Aesthetic* that,

the concept of alteration and, with it, the concept of motion (as alteration of place), is only possible through and in the representation of time – that if this representation [i.e. the representation of time] were not *a priori* (inner) intuition, then no concept, whatever it might be, could make comprehensible the possibility of an alteration, i.e., of a combination of contradictorily opposed predicates . . . in one and the same object. (A31/B48)

The concept of alteration, then, is *a priori*. It is not, however, a *pure* concept because it is impossible to think of an alteration without thinking of its taking place in time. The notion of alteration does involve pure categories, but it also involves reference to our representation of time, and this representation of time, although it is *a priori*, is not a product of our intellect but a product of our faculty of intuition. In the concept of alteration pure concepts of the understanding (categories) are “mixed with something foreign” to them, namely our intuition of time. Only concepts which are products of the pure understanding, then, are, in the strict sense, pure concepts.⁴³

⁴² The schematized concept of cause is, then, on my interpretation not a pure concept. Although, it is a pure schema.

⁴³ Thus the unschematized categories are pure, the schematized categories *a priori*, but not pure concepts.

This understanding of purity lies behind Kant's distinction between pure philosophy (or metaphysics), on the one hand, and mathematics (and natural science) on the other.⁴⁴ As

Hartman (1974) explains in his introduction to Kant's *Logic*,

the pure part of natural cognition must show what reason can accomplish by itself and where it needs the assistance of principles of experience. **Pure rational cognition based on mere concepts is called pure philosophy or metaphysics**; whereas rational cognition based on the construction of concepts, through representation of an object in an intuition *a priori*, is called mathematics. (p.lxxix – my emphasis)

According to Hartman, then, for Kant mathematics is *a priori*, but not pure. My claim is that, for Kant, our examination and understanding of the idea of a good will must be based upon pure philosophy understood in this sense.⁴⁵

To understand the distinction between concepts of pure reason and merely *a priori* concepts, then, one can consider the difference between our ideas of pure reason (such as God, or spirit) and the categories, which are pure concepts, and our (mathematical) concepts of numbers and shapes, which are *a priori* but not pure concepts. Numbers and shapes, then, not being empirical concepts, are thinkable *a priori*. They are not concepts of the *pure* intellect, however, because thinking them, according to Kant, also requires some input from the faculty of intuition. For numbers and shapes must be constructed *a priori* in pure intuition (in our pure intuitions of space and time). In other words, our (mathematical) concepts of numbers and shapes (and our concept of alteration) are not pure because they require the co-operation of two faculties: the faculty of intuition and

⁴⁴ See Kant's *Preface to the Metaphysical Foundations of Natural Science* for this distinction.

⁴⁵ Mathematics involves the *construction* of concepts (in intuition), *pure* philosophy, in contrast, starts with analysis, and Kant claims that moral philosophy is 90% analysis.

the faculty of understanding/reason. This can be expressed without making any reference to the notion of distinct faculties, in terms of the difference in content between these two types of concept. Concepts of number and shape are not pure because they have both a categorical (conceptual) and a spatio-temporal content. In contrast, ideas of pure reason, being pure, are thinkable without any input from the faculty of intuition and as such have no spatio-temporal content (or properties). They are, Kant believes, constructed purely from *a priori* concepts in pure thought, whereas our concepts of shapes are constructed in (the *a priori* intuition of) space, while numbers are constructed in (the *a priori* intuition of) time.⁴⁶ The *pure* idea of a world, for example, can have no spatio-temporal content (and as a consequence can have no spatio-temporal qualities or properties).

A *pure* idea, then, is one that originates purely from the intellect, with no input from the faculty of intuition. Thus, if our moral ideal is to be pure, it must not only be *a priori* but it must also have no intuitive content, that is, it must have no content provided by the faculty of intuition. In other words, Kant maintains that our moral ideal is ‘intelligible’; it is the idea of an individual conceived of in a non-spatio-temporal way.

Our moral ideal is the pure idea of a member of a world

Our moral ideal, then, must be a pure idea. But, Kant believes, we possess a number of pure ideas, and so the questions must be asked: which of our pure ideas can function as

⁴⁶ Or, to be more precise, our concepts of numbers and shapes involve an awareness of how they are *constructible* in time or space. They are necessarily schematized. We cannot have the unschematized concept of a number or a shape. See the *Schematism*, A141/B180.

our moral ideal? Kant argues that to be moral is to choose to be a member of a moral world or a kingdom of ends. And he identifies the idea of a moral world with the idea of an intelligible world.⁴⁷ For Kant, then, the choice to be moral is the choice to be ‘a member of an intelligible world’.

In talking of an ‘intelligible world’, at least in the ethical context, Kant is not referring to some other, ontologically distinct world. Instead, when he talks of the idea of an intelligible world all he really means is our pure idea of a world, that is our idea of a world insofar as it is an object of the pure intellect.⁴⁸ In many contexts in which Kant talks of “the intelligible world”, then, we could replace this expression with “our pure idea of a world”, and in this dissertation I will use these expressions interchangeably. On this interpretation the word ‘idea’ is redundant in the expression ‘our idea of an intelligible world’. I will, however, occasionally use this expression, just to remind the reader that when Kant is talking of ‘the intelligible world’ he is normally just talking about our *idea* of a world. Kant makes this clear in the *Critique of Pure Reason*, where he explains that,

The *mundus intelligibilis* is nothing but the concept of a world in general, abstracting from all [spatio-temporal] conditions of intuiting it. (A433/B461 – my addition in square brackets)

Here Kant makes it clear that when he talks of the intelligible world, or *mundus intelligibilis*, all he means is our pure concept of a world. He also makes this clear in his

⁴⁷ Thus, in the *Critique of Pure Reason* he can write: “an intelligible, i.e., a **moral** world” (A811/B839).

⁴⁸ In the expression the “idea of the intelligible world”, grammatically, “intelligible” is modifying “world”, but Kant is writing clumsily because it should really be modifying “idea”, stressing that the idea is a pure idea, that is, a product of the pure intellect. “Pure” and “intelligible”, in this context, amount to the same thing.

metaphysics lectures, where he argues that, “there is also an intelligible world that exists merely in my understanding apart from all possible experience; but in this, space and time are out of the question” (29:858).⁴⁹ The intelligible world, then, is merely an object of pure thought, or an idea; it is the idea of a world that is not spatio-temporal and which has no objective reality [at least from the theoretical perspective]. In other words, this idea cannot be an object of our faculty of intuition and as a result cannot be an object of experience. However, although such a world cannot be an object of (our faculty of) intuition, it can be an object of choice, or what Kant calls our faculty of desire. For he believes that although it is impossible for us to experience a world corresponding to our pure idea of a world⁵⁰, it is possible for us to choose to be a member of such a world. He does not, however, believe that there is any theoretical justification for believing that such a world, corresponding to our idea, actually exists, or even that such a world is really possible. Our pure idea of a world, then, contains no contradictions, and so it is conceivable that a world could actually exist corresponding to this idea. Given the conceivability of such a world there is nothing irrational in making it our life project to be member of such a world.

Our idea of an intelligible world is the idea of a *community* of individuals in interaction⁵¹

⁴⁹ Elsewhere he argues that, “a foreigner called it fantasy to speak of the **intelligible world** <*mundo intelligibili*>. But this is just the opposite, for **one understands by it not another world, but rather this world as I think of it through the understanding**” (*Metaphysik Mrongovius*, 29:850 – my emphasis).

⁵⁰ Or even for us to experience something corresponding to our pure idea of a of a member of a world.

⁵¹ This account seems to be at odds with certain claims Kant makes in the first *Critique*. In particular, in the first *Critique*, Kant distinguishes between three ideas of pure reason – namely the ideas of the soul, the world and God. Each of these is derived from one of the three categories of relation. And in the first *Critique* the idea of the world is not derived from the category of community but from that of the second

Kant believes that our idea of a world is, by definition, the idea of a whole. Kant states this explicitly and frequently in his metaphysics lectures. For example in his lectures from 1792-3 he argues that,

A multitude of substances without connection makes no world. One must thus not define world: the universe of substances, but rather the whole of them. (*Metaphysik Dohna*, 28:657)

And he makes it clear that he believes that we can only think of a whole of substances if we think of the substances as connected and interacting.⁵² Kant continues by introducing the distinction between the form and matter of a world and arguing that while the “*material of the world are substances*”, the formal element

category of relation and thus here the idea of a world is the idea of a series and not the idea of a community. I suggest that in the first *Critique* he is giving an account of our idea of the phenomenal world, and not an account of our idea of a noumenal world. It should be clear, however, that when he talks of the intelligible world in his metaphysical writings and in his ethical work he has something more in mind than the idea of a series. I suggest that Kant ultimately believes that there is only one idea of pure reason – but that this idea is complex. Specifically I suggest that Kant believes that we have the complex idea of ‘a world (or community) of individual substances created by God’. I suggest that none of the component ideas, distinguished in the first *Critique*, can ultimately be thought in separation from one another. And hence, that each idea (for example the idea of a ‘world’ or ‘community’) involves all of the categories. I have discussed this in slightly more detail in chapter one. And I believe that the account offered here is ultimately compatible with Kant’s account in the first *Critique*. However a defense of this belief, and in particular an account of how the three ideas of pure reason identified in the first *Critique* are derived from the three categories of relation, is beyond the scope of my dissertation. I will merely suggest that in the first *Critique* Kant is interested merely in examining three aspects of this complex idea. Thinking of the three ideas of pure reason identified in the first *Critique* as aspects of one unified complex idea helps us understand the *unity* of reason.

⁵² Kant’s claim that a multitude of substances can only constitute a world insofar as they interact should be read as a denial of the Leibnizian claim that isolated monads can constitute a world, or even a possible world, on the basis of the (pre-established) harmony between them. At this stage I have suggested that individuals can be thought of as being members of a world only insofar as they are in interaction. In Chapter 6, when I examine Kant’s analysis of action, I will argue that individuals can only interact if they are already members of a community, and so one cannot explain how it is possible for individuals to be members of a unified community on the basis of the fact that they interact. So Kant’s account of how individuals can be thought of as members of a unified community cannot appeal to their interaction. In Chapter 7, when I examine Kant’s account of our idea of community I will explain how his analysis of this idea is not based upon presupposing the interaction of the individuals that constitute a community.

is the real connection <nexus realis> of these substances. Real connection is reciprocal influence (acting and suffering). . . a multitude of substances without connection makes no world. (Metaphysik Dohna, 28:657)

The idea of a world, then, is the idea of a whole of (individual) substances, and we can only think of a whole consisting of individuals if we think of the whole as a community and the individuals that constitute the whole as really interacting. As a consequence of this definition of a world, the idea of a member of a world is, by definition, the idea of ‘an individual in real interaction with other individuals’.

This identification of the idea of a world with the idea of a community is not an off the cuff remark. Because this identification is so important for my thesis I will provide some more textual evidence for my interpretation.

In his lectures from the early 1780s Kant once again distinguishes between the idea of a world and that of a mere multitude and argues that,

a great multitude of isolated substances would not constitute a world (isolated substances are only the stuff for a world), because they would not constitute a whole, but rather each of them would be entirely alone and without any **community** with the others. (*Metaphysik Mrongovius*, 29:853 – my emphasis)

Substances, then, can only constitute a world if they are in community with one another.

Kant's belief, that our idea of a world is, by definition, the idea of a community of substances (or individuals), can be traced back to his pre-critical period. Thus, in his metaphysics lectures from the mid-1770s he claims that,

The aggregation of the substances in which there is no community still does not constitute a world. Reciprocal determination, the form of the world as a composite <compositi>, rests upon interaction <commercio>. **If we thought substances without real connection <absque nexu reali> and without interaction <commercium>, where every substance would be in and for itself and they would have no community with one another, then that would indeed be a multitude <multitudo>, but still not a world.** // Thus the connection <nexus> of substances that stand in interaction <commercio> is the essential condition of the world. (*Metaphysik* L2, 28:196 – my emphasis)

Once again Kant makes it quite clear that the distinction between a world and a mere multitude is that for a set of individuals to constitute a world there must be some ‘real connection’ between them, that they must interact and constitute a community. Similar passages are not hard to find.⁵³ The idea of an intelligible world, then, is the pure idea of a community of substances. And, as we have seen in chapter one, I believe that the idea of a kingdom of ends is merely the idea of a world, understood in this sense as a community of substances in real interaction.

To choose to be moral, then, is to choose to be a member of a community. In some sense, then, Kant’s ethics should be understood as ‘communitarian’. This is not to say that to be moral is to choose to be a member of some actual community; rather, to be moral is to choose to be a member of an ideal community. In saying this I agree with O’Neill (1989), who argues that, “the Categorical Imperative states essential requirements for a possible community (not an actual community) of separate, free and rational beings” (p.44). O’Neill, however, does not explicitly draw the conclusion that this implies that an explanation of the possibility of morality involves an explanation of how we can actually

⁵³ See for example: 28:581-2, 28:45, 29:851-2, 29:868, & 29:1006-7.

have an *a priori* idea of such a community. For Kant believes that our idea of ‘community’ is an *a priori* and pure idea and we know by examining this idea *a priori* how we would have to act if we decide we desire to be a member of a community. It is instructive, at this point, to compare Kant with Hume. Hume is an empiricist, believing that all of our ideas must have an empirical origin. Hume’s empiricism informs his ethics and political philosophy, for he believes that our idea or conception of society must be based upon experience. As a result of this he argues, in the *Treatise*, that the idea of a pre-social duty to enter into society is non-sensical, because we cannot actually acquire the idea of society until we have actually experienced a society.⁵⁴ Hume’s extreme empiricist assumptions, then, form the basis for his rejection of Locke’s political philosophy, for Locke argues that, even in the state of nature, man would be capable of recognizing a reason to enter into society, which implies that man must be capable of possessing the idea of society before he actually experiences one. Kant sides with Locke on this issue, believing that we have a duty to enter into society. As a consequence he believes that our idea of society must be *a priori*.⁵⁵ Unlike Locke, however, Kant believes that he can provide a detailed account of what is involved in this *a priori* idea of society or community, for he believes that we can trace this idea back to the category of

⁵⁴ Thus, Hume (1978) argues, presumably against Locke, that, “in order to form society, ‘tis requisite not only that it be advantageous, but also that men be sensible of its advantages; and ‘tis impossible in their wild uncultivated state, that by study and reflection alone, they should ever be able to attain this knowledge” (p.486). This is the reason why Hume claims that justice is an artificial virtue. The reason why we cannot, in the state of nature, have knowledge that society would be beneficial is because in the state of nature we do not have the idea of society. I suggest the reason for this is Hume’s belief that, “all beings in the universe, consider’d in themselves, appear entirely loose and independent of each other. ‘Tis only by experience that we learn their influence and connexion” (p.466).

⁵⁵ Unlike Locke, however, Kant believes that he has the philosophical resources, lacking to the empiricist, to explain the nature and ‘genesis’ of this idea of society. For the idea of community is based on the category of community, which in turn is derived from the disjunctive form of judgment. In Chapter 6, I will explain, in more detail, the *a priori* nature of the idea of community, its basis in the table of judgments and its relationship to the idea of autonomy.

community and the corresponding disjunctive form of judgment. In chapter eight I will examine Kant's account of the origin of this idea in more detail.

The idea of a member of a community (or world) is the idea of an autonomous agent

Reflecting on what is involved in our pure idea of a world, then, we realize that in order to be a member of a world an individual must have certain characteristics. In particular, Kant believes that our pure idea of a community (or world) can only be the idea of a community of autonomous agents, each of whom 'gives' or legislates the laws of the community. As a consequence of this, the pure idea of a *member* of a world is the idea of an autonomous agent. The reason for this is that Kant believes that a world is essentially unified, for, as we have seen, it is this unity that distinguishes the idea of a world from that of a mere multitude. In addition, he believes that a world can only be unified or 'held together' by laws, with such laws being understood to be intelligible relations. So the idea of a world is the idea of a multitude of individuals unified by laws. Now, if the unity of a world is to be 'intrinsic' to the world, rather than merely existing in the mind of some ideal observer observing the world, then the members of the world must be responsible for the unity of the world, and Kant believes that this is only possible if each individual member of the world is the source of, or 'the giver of', the laws that provide the world with its unity. In other words, each member of a world must be autonomous. Thus, our pure idea of a world is the idea of a multitude of autonomous agents.

If we decide to make it our fundamental project to be a member of an intelligible world we must behave in a way that makes this possible. In particular we must choose to be autonomous, to be legislators, giving and abiding by the laws of such a potential world. If we act in such a way we have no theoretical grounds for assuming that we will actually become members of such a world, for we have no theoretical grounds for assuming that there is actually, or even could be, such an intelligible world corresponding to our idea. The pure idea of a world, however, does not contain any contradictions, so there is nothing irrational in hoping to become a member of such a world. This involves hoping that there is an intelligible world corresponding to our pure idea of a world which actually exists, and that if we act suitably we can become members of it. Or perhaps, and I think that this is perhaps Kant's considered opinion, we can hope that even if such a world does not actually exist, we can actually bring it into being through our willing.

It should be clear from what I have said above that I am advocating a particular conception of autonomy. To be autonomous, on this interpretation, is not merely to 'give law to oneself', but to give a particular type of law to oneself. To be autonomous is to give laws that bind both oneself and others. One could, however, imagine a solipsistic egoist who wants to give some unity to his life and so chooses to act only on certain principles or laws. Such an egoist attempts to give law to himself, but the only law he attempts to submit himself to is an intra-personal law. Kant suggests that Wolff can be thought of as advocating such an ethical principle, for the principle of perfection demands that we unify our desires and inclinations, but not necessarily in a way that makes them compatible with the desires and inclinations of others. Nietzsche also, at times, seems to

advocate a sort of egoistic, solipsistic autonomy.⁵⁶ Such an individual, who attempts to give unity to her inclinations by subjecting them to some personal law, might be thought of as taking Leibniz's conception of a monad as her moral ideal. The Leibnizian believes that the only type of finite individual we can conceive of is a solitary individual. Such an individual is essentially active and its activity is that of having representations. A Leibnizian monad, then, can be thought of as a series or stream of representations. This stream, however, is essentially unified. What unifies the representations of an individual is that they are subject to a law, and Leibniz calls this law, which provides the representations of an individual their unity, the 'law of the individual' or the 'law of the series'. The law of the series can be thought of the source of the series of representations and, Leibniz believes, accounts for the unity of the individual. Such a law, however, is to be regarded purely as an internal law. An egoist, then, who takes such a conception of an individual as his ideal would try to unify his representations (or desires) but thinks that this can be done purely by reference to some law purely internal to the himself, perhaps the 'law of his genius'. The 'rational' solipsist (or a rational hedonist), then, can be thought of as attempting to subject herself to purely **intra-personal** laws, which make no reference to other individuals. Such an individual may claim that they are striving to be autonomous, claiming that they subject themselves to their own laws, or perhaps that they subject themselves to the law of their own genius. An autonomous agent in Kant's sense, however, legislates and subjects himself to **inter-personal** laws, that is, to laws of a

⁵⁶ Thus, in *Thus Spoke Zarathustra* (Nietzsche, 1969), in the section *Of the Way of the Creator*, Zarathustra asks: "Do you call yourself free? I want to hear your ruling idea, and not that you have escaped from a yoke. . . Free from what? Zarathustra does not care about that! But your eye should clearly tell me: free *for* what? Can you furnish yourself with your own good and evil and hang up your own will above yourself as a law? Can you be judge of yourself and avenger of your law?" (p.89). Nietzsche seems to suggest that we must be creative in the sense of creating our own, individual intra-personal law.

possible (ideal) community. To be autonomous, in the Kantian sense, then, is not merely to legislate for oneself but for a (potential) community.⁵⁷

Deciding to be moral, then, is a bit like deciding to join a club. One can imagine the idea of an ideal club and decide that the most important thing in life is to be a member of such a club. The idea of such a club, to be an adequate idea, will include, in general, a conception of what the membership criteria are. Now we don't know whether such a club exists or not. And there is nothing I can do alone to bring this club into existence – for the existence of the club, even if such a club is possible, depends upon others freely choosing to become members (for a club of one is not a club), and this is not something I have, or could have, any control over. If, however, I decide that being a member of this club is the only important thing in my life, then I will make satisfying the membership criteria the primary concern of my life, and must just hope that there are others who (a) know the rules of the club (i.e. that there are others who are conscious of the moral law – the **same** moral law I am conscious of) and (b) decide to become members.

To conclude: A moral idealist is someone who believes that at the heart of our ethical judgments and practices lies the pure *a priori* idea of a particular kind of individual. My claim is that Kant is an ethical idealist and that such an 'ideal', which he often calls 'the idea of a good will', lies at the heart of Kant's ethics. The main aim of this dissertation will be to examine what is involved in this ideal. I will argue that the moral ideal of a 'good will' is to be identified with the 'cosmological idea' of 'a member of an intelligible

⁵⁷ In claiming this I am disagreeing with Schönfeld (2000), who argues that, "Leibniz's pre-established harmony permits the autonomy of souls" (p.141).

world', and that we can understand many of Kant's ethical claims by looking at what he has to say about the idea of 'an intelligible world' in his metaphysical works. The most fruitful sources of information on how Kant conceived of an intelligible world are his pre-critical writings and his *Lectures on Metaphysics*, and these sources will be referred to frequently in this thesis.

(2b) Kant's Phenomenal Solipsism

In claiming that Kant advocates an 'ethics of interaction' (or perhaps an 'ethics of relations') I mean that Kant's moral ideal is the idea of being an individual in interaction with other individuals and that, as a consequence, to be moral is to choose to interact with others. Indeed, one could say that, for Kant, to be moral is to choose that others exist. The reason for saying this is because Kant believes that the existence of (other) individuals is not something that is given to us in experience, rather we merely have a pure *a priori* idea of the other, and nothing given in experience can be adequate to this idea.⁵⁸ To understand what this might mean it is necessary to understand Kant's attitude towards solipsism and 'the problem of other minds'.

Kant believes there is and can be no theoretical answer to the solipsist. We have no way of *knowing* or even reasonably inferring that other individuals actually exist.

⁵⁸ Indeed, we are not aware of our own existence as an individual. We have the idea of individuality, but do not, and cannot, have the experience of an individual. Not only do I have to choose that others exist as individuals, I am faced with the choice of whether I want to make it my project to exist as an individual. These two choices, however, go hand in hand.

Theoretically there is no escape from solipsism, or what Kant calls ‘egoism’.⁵⁹ Phenomenal experience provides us with no evidence for the existence of other individuals – indeed it provides us with no evidence that we ourselves are really individuals. This does not, however, imply that we must become solipsists or egoists. For the ‘problem of other minds’ is not a theoretical problem but a moral one. We are free, Kant believes, to treat the world around us as a world of objects, regarding the living bodies around us as complicated machines merely to be used. We can, however, choose to regard other living human beings as individual persons by attempting to interact with them intelligibly; this is what we do when we choose to listen to our conscience. We cannot *know* that the human beings we encounter around us really are rational individuals, for we do not experience them as individuals, but we can choose to *think of them* as individuals without contradiction.

A number of Kant scholars, most notably Strawson, Bird and Walsh, who have attempted to give a Wittgensteinian interpretation of Kant, would disagree with this solipsistic reading of Kant’s theoretical philosophy. For, on their interpretations, the conditions for the possibility of judgment that Kant enumerates in the *Critique of Pure Reason* should be understood as conditions for the possibility of intersubjective agreement. On such an interpretation, Kant’s whole theoretical framework presupposes the existence of other minds.

⁵⁹ Kant makes it clear that he identifies egoism with solipsism in his metaphysics lectures. For example in his lectures from 1792-3 he claims that, “the Egoist is he who holds himself, as thinking being, as the only worldly being” (*Metaphysik Dohna*, 28:663).

In his review of Walsh's *Kant's Criticism of Metaphysics*, Guyer (1977) rejects such Wittgensteinian interpretations of Kant, arguing that, "intersubjectivity plays no role at all in the main argument of the *Critique of Pure Reason*" (p.268). Guyer argues that,

Such interpretations give both the concept and the possibility of other persons a force which they are not meant to have in Kant's Transcendental Deduction. Instead, Kant's argument adopts the standpoint of methodological solipsism in order to construct the strongest possible answer to an empiricist skepticism. (p.266)

Guyer is right to stress that Kant's theoretical arguments in the *Critique of Pure Reason* make no appeal to the notion of intersubjectivity or to inter-personal agreement. However, I do not think Guyer goes far enough here, for Kant does not merely 'adopt the attitude of methodological solipsism'.⁶⁰ Instead he is committed to what I shall call phenomenal solipsism (or egoism). In other words, Kant believes that the world actually appears to us solipsistically. The world we experience is experienced, Kant believes, as a world in which there are no true individuals. We do not, then, experience other human beings as individuals or persons; rather, it is up to us, *from the practical perspective*, to choose whether or not recognize other human beings as individuals existing independently of us and to treat them as persons. Kant, then, is not merely committed to the position that from the theoretical perspective there is no answer to the solipsist, but

⁶⁰ I should point out at this juncture that, unlike Guyer, I do not see Kant's primary goal as attempting to provide a answer to (empiricist) skepticism. Instead I regard him as engaged in a type of phenomenology, attempting to describe the nature and structure of our experience (in such a way as to leave morality possible). I do not, then, think he is primarily concerned with justifying our judgements, but with explaining how we are able to make them. As a result, I do not regard his solipsism in the *Critique of Pure Reason* as 'methodological'. Although Kant himself argues for the primacy of the practical over the theoretical I do not believe he goes far enough. He seems to believe that the account of experience he offers in the first *Critique* is true. For me it is enough that it is a plausible story, and one that leaves room for the possibility of morality.

also to the stronger position that if the phenomenal world were all there were then the solipsist would be right.

Kant's remarks on Spinoza in his published writings makes his attitude clear. For Kant repeatedly maintains that if things in themselves really did exist in space and time as we experience them, then Spinoza would be right in believing that God is the only real substance or individual. And Kant believes that to claim that God is the only true individual or substance is equivalent to claiming that God is the only real being. Thus, in the *Critique of Practical Reason*, for example, Kant explains that,

If this ideality of time and space is not adopted, nothing remains but Spinozism, in which space and time are essential determinations of the original being itself, while the things dependent upon it (ourselves therefore included) are not substances [i.e. simple, individuals] but merely accidents inhering in it; (5:102 – my addition in square brackets)⁶¹

This is why Kant believes that the 'realist' believes "that there is nothing present outside him, but rather everything that we see is mere illusion". For we must remember that Kant is working in the Leibnizian tradition. In this tradition all beings must ultimately be simple and indivisible, or composed of simples. As Leibniz (1989) famously claims, "what is not truly *a* being is not truly a *being*. It has always been thought that one and being are reciprocal terms" (p.85). The (transcendental) realist claims that the objects of experience, or bodies, are the only things that are 'real'; what you see is what you get. Bodies, however, are not simple; existing in space, they are, Kant believes, infinitely

⁶¹ Kant also equates transcendental realism with Spinozism on a number of occasions in his metaphysics lectures, claiming, for example, that, "if we consider space as real, we assume Spinoza's system" (*Metaphysik Dohna*, 28:666). See also: 28:567, 29:1008-9 & 29:977-8.

divisible.⁶² So to claim that bodies are all that is real is, in effect, to deny that anything really exists independently of us. Thus, if we accept Kant's Leibnizian assumption that everything real must be simple, the realist, in affirming that (apart from ourselves) only bodies possess reality, is in fact committing himself to what Kant calls empirical idealism, for he is implicitly committed to the position that there is nothing real external to us.

Realism and idealism

Kant is famous for claiming to be an empirical realist but a transcendental idealist. And he believes that a commitment to what he calls 'transcendental realism' (and Kant would class both contemporary naturalism and physicalism as forms of transcendental realism) is equivalent to being an empirical egoist or solipsist. This is what he means when he claims in the *Critique of Pure Reason* that, "transcendental realism. . . finds itself required to give way to empirical idealism" (A371), for empirical idealism is equivalent to solipsism or egoism. This identification of realism with solipsism may sound surprising, because there is a tendency to understand the terms 'realism' and 'idealism' anachronistically. To understand what Kant means by claiming to be a transcendental idealist but an empirical realist, we need to understand how the terms 'idealism' and 'realism' were used in 18th century German metaphysics; 'realism' and 'idealism' were understood to be alternatives to (Cartesian) dualism. Thus, in 18th century metaphysics

⁶² As he argues in his lectures from 1792-3: "[I]n the world of appearances, there are no simple parts. Only in the intelligible world, noumenal world is of monads, but we do not at all cognize it" (Metaphysik Dohna, 28:663-4).

textbooks, questions about idealism were discussed as a part of ‘Rational Psychology’, and the standard distinctions were between egoism, dualism and idealism. Understood in this context, idealism, dualism and egoism are metaphysical positions concerning what types of substances there are. Kant will argue that ‘dualism’ is untenable and that ‘idealism’ and ‘realism’, at least insofar as they were traditionally understood, collapse into one another. Kant himself believes that he has a far more tenable position, which he calls transcendental, or formal, idealism.

‘Idealism’ can be understood in two senses. On the one hand idealism can mean the denial of the reality of the world around us. On the other hand, it can mean the claim that (only) ‘spirits’ or ‘minds’ really exist independently of us. I will call the first conception of idealism ‘idealism in the negative sense’ and the second ‘idealism in the positive sense’.

Baumgarten defines ‘idealism’ in his *Metaphysics*, the textbook that Kant used for his metaphysics lectures for over 30 years, in the following terms: “He who admits only spirits in this world is an IDEALIST.” (#402) Kant follows this definition in his own lectures but then identifies the alternative to idealism, so understood, not as realism but as egoism.⁶³ Given the fact that, traditionally, realism was considered to be the alternative to idealism, labeling the alternative to idealism ‘egoism’ provides strong evidence that Kant believes that the (traditional) realist must be an egoist or solipsist. Kant, then, contrasts

⁶³ In fact, following Baumgarten, Kant distinguishes between egoism, idealism and dualism. Claiming that, “the dualist believes in thinking beings and bodies outside him.” (29:928) However, Kant rejects dualism out of hand arguing that, “in general it [Dualism] is so absurd that it may well never occur to affirm this error seriously, even if it were irrefutable as well” (29:928).

the **egoist** who maintains that, “there is nothing present outside him, but rather everything that we see is mere illusion” (29:927), with the **idealist** (and here Kant is clearly talking of idealism in the positive sense), and explains that, “idealism is when one imagines that outside oneself thinking beings are indeed present, but not bodies” (29:928).⁶⁴ Kant continues by maintaining that the egoist cannot be refuted, at least not by our experience.⁶⁵

In his ethical writings Kant advocates something like idealism in the positive sense, for to be virtuous is to choose to (or at least to strive to) regard ourselves and other human beings as individuals. It is to choose to interact with other individuals, with other ‘spirits’, and not just in some fantasy world, but in the physical world we experience. However, we do not, and cannot, experience physical objects (including the physical human beings around us) as individuals. Indeed, we may well ask: How is it possible for a physical object to be an individual agent (or ‘spirit’)? For physical objects are necessarily spatio-

⁶⁴ Of course, in claiming to be a ‘transcendental’ idealist, Kant ultimately wants to reject both realism and idealism (or perhaps we could say he wishes to accept them both, merely denying their negative claims). The purpose of this discussion, however, is merely to show that Kant follows Baumgarten in claiming that the alternative to idealism is egoism, which implies that Kant believes that ‘realism’ and ‘egoism’ are interchangeable terms. I believe that, following Leibniz, Kant believes that everything that exists must be simple. The realist, then, is a solipsist because he believes that what we see is what we get, and we never experience anything simple. The idealist, on the other hand and rather paradoxically for the contemporary reader, actually affirms the existence of true individuals independent of us. Kant, being a transcendental, or formal, idealist will claim that the lack of simplicity in the objects of experience is not due to the objects themselves but is due to our form of intuition. Thus Kant explains in his lectures on metaphysics from the 1790s that, “Whoever maintains and assumes ideality with respect to the form, that space and time are not properties, but rather are only subjective conditions of our intuition, he is a transcendental idealist” (28:773). Kant explains the benefit of adopting transcendental idealism in the preceding sentence. Here Kant claims that, “[if e]xtension has its ground in my representation; the thing itself can be simple”.

Transcendental idealism, then, provides a way of explaining how we can still think of objects of sense (and in particular the other human beings we experience) as simple, even though we can only experience them as infinitely divisible. Transcendental idealism, then, allows us to assume that the lack of simplicity in the objects of experience is due to our manner of experiencing things and not due to the things themselves.

⁶⁵ “I cannot refute the egoist by experience, for this instructs us immediately only of our own existence” (29:927).

temporal, whereas spirits are by definition non-spatio-temporal. To act morally involves regarding (and treating) other bodies as autonomous individuals.⁶⁶ The problem the ethicist faces is that all objects of experience (bodies, including human bodies) are experienced as non-individual.⁶⁷ One solution would be to adopt a dualist position and argue that bodies and minds (spirits) are two distinct types of things. Kant, however does not want to take this route for it would suggest that we don't really have any real ethical duties towards the human beings (bodies) around us, but, at most, only to the minds they are somehow connected to. Kant rejects this route because he believes that: (a) we can only have ethical commitments towards individuals but, (b) we have ethical commitments towards the human beings (bodies) around us.⁶⁸ It is only possible to hold both of these commitments if it is possible for us to *think* of the human beings (bodies) around us as individuals without contradiction. The problem, however, is to explain how it is possible to think of bodies as (simple) individuals given that they are experienced as essentially spatio-temporal, and hence as non-simple.

So far in this discussion I have couched this discussion in terms of the necessary individuality of an autonomous individual. Some commentators may not be sympathetic towards my stress on the importance of the notion of individuality in Kant's ethics. An

⁶⁶ It is not clear to me whether Kant thinks there is any distinction between the choice to really *recognize* other human beings as autonomous and the choice to *treat* other human beings as autonomous. I suspect that Kant's considered opinion is that we can only treat others human beings as not autonomous insofar as we choose not to recognize them as autonomous. This is a consequence of his belief, discussed in chapter one, that the commands of morality are like Sirens voices, and in so far as we choose to listen to the commands of morality they are irresistible.

⁶⁷ This is a rather clumsy expression, but it is important to stress that Kant does not merely make the claim that (human) bodies are not experienced as individuals, but also makes a stronger claim, namely, that bodies as experienced are experienced as not being individuals.

⁶⁸ Kant's commitment to (b) is a result of his respect for common sense, pre-philosophical, ethical commitments. For (b) is a central commitment of common sense morality.

analogous argument could be made in terms of the necessarily unconditioned nature of our idea of an autonomous individual.⁶⁹ Kant believes that we can only have moral relations with unconditioned (free) beings, but that we experience every body as essentially conditioned, and, thus, he must be able to explain how it is possible for us to think of bodies as unconditioned without contradiction.

Kant's solution to the mind-body problem - transcendental (formal) idealism

In the *Sixth Meditation*, Descartes famously argues that mind and body must be distinct kinds of substance because (i) mind is essentially simple (indivisible) whereas (ii) body is essentially divisible, and (iii) one kind of substance cannot be both essentially divisible and essentially indivisible (simple). Kant accepts premises (i) and (ii), but he cannot accept the conclusion, and so rejects premise (iii).⁷⁰ The rejection of premise (iii) is transcendental idealism. But how can we possibly reject premise (iii)?

Kant's answer will be to claim that although (a) the objects of experience are real, (b) their (infinitely) divisible form is ideal or merely apparent.⁷¹ This offers a solution to the mind/spirit problem⁷², for it allows us to explain how we can *think* of the bodies we experience as simple (minds or spirits) without contradiction, even though it is a

⁶⁹ Although, here, a critic of Kant could object to Kant's move from the claim that our *idea* of an autonomous being must (as an idea) be unconditioned to the claim that our idea of an autonomous being must be the idea of an unconditioned *being*.

⁷⁰ Although, for Kant, the problem isn't the *mind*-body problem, but the *spirit*-body problem.

⁷¹ (a) is Kant's empirical realism, (b) his transcendental, or formal, idealism.

⁷² In his lectures Kant customarily introduces transcendental idealism in the context of his discussion the mind-body problem and rational psychology. Similarly, in the *Critique of Pure Reason*, Kant discusses transcendental idealism in the *Paralogisms*, a section that corresponds to rational psychology in traditional metaphysics textbooks.

condition of our experience that we experience them as essentially non-simple.⁷³ Transcendental idealism, then, leaves conceptual space for us to *think* of the objects we experience as simple even though they necessarily appear to us as non-simple, and this allows us to accept the reality of both bodies and spirits ('minds') without accepting dualism. It does this by postulating a distinction between the form and the matter of the objects of experience and affirming the reality of the matter of the objects of experience while denying the reality of their spatio-temporal form. The spatio-temporal form of these objects is ideal and subjective. This allows room for us to *think* of (at least some) objects of experience as simple (as spirits) without contradiction, and this is necessary if we are to have moral relations towards human bodies. If we reject transcendental idealism, and, following the transcendental realist, affirm the reality of both the form and matter of the objects of experience, we must maintain that the world is in itself spatio-temporal as it appears to us. This would imply that there is nothing simple independent of us. And, if we assume with Leibniz that what is not truly *a* being is not truly a *being*, to maintain there is nothing simple independent of us is to claim that nothing really exists independently of us. Thus, given the untenability of dualism, if we accept the Leibnizian assumption, the transcendental realist can only be an egoist or solipsist. Transcendental idealism, on the other hand, in claiming that the lack of simplicity we experience in the world is merely due to our form of intuition, leaves room for the logical possibility that some of the bodies that we experience as non-simple are in fact, in themselves, simple

⁷³ Similarly it also explains how we can think of other people as unconditioned even though we experience them as conditioned.

individuals.⁷⁴ Whether we choose to regard some of the bodies around us as true individuals is not, however, a theoretical matter, but a practical one.

In our daily lives, then, we encounter many bodies. We do not experience them as simple. We do not experience them as individuals. This is true even of the human beings we encounter. We know that human beings can die and decompose. They can be cut up and burned. They can all disappear into crowds. This is not, for Kant, a moral claim, but a claim about the nature of our experience. It is just a fact about our experience that we experience nothing as simple and truly individual. For, Kant believes, the idea of an individual is the idea of something simple.⁷⁵

Conclusion

Our moral ideal is the idea of a member of an intelligible world (or community), and this is the idea of an autonomous individual. Individuality is, however, intelligible. An

⁷⁴ This interpretation of Kant is diametrically opposed to Schopenhauer's, for Schopenhauer argues that, for Kant, individuals exist only phenomenally, whereas the intelligible realm is one. Thus Hollingdale writes (in Schopenhauer, 1970) that, according to Schopenhauer, "since Kant had proved that space and time, the spatio-temporal fragmentation of the world, belonged as forms of perception only to the plane of phenomena, the noumenal plane, the 'real world' of the thing in itself, must be one and indivisible. Consequently the will in the stone and the will in me is the same will" (p.21). However, Kant is driven to transcendental idealism because he wants to avoid Spinozism, for "Those who assume space as a matter in itself or as a constitution of things in themselves, are required to be Spinozists, i.e., they assume the world to be the summation of the determinations of a united necessary substance, thus only one substance" (29:1009). Transcendental Idealism, however, allows us to *avoid* making this assumption.

⁷⁵ I am focusing on 'simplicity' here for the sake of argument. Kant believes that 'spirits' must be autonomous individuals, and as individuals they must be simple, in addition, however, a part of being autonomous is the idea of being a cause which is not an effect. Kant also believes that we cannot experience anything as a cause which is not an effect. Indeed it is a feature of the temporal nature of our experience that everything we experience must be experienced as an effect. Transcendental idealism allows us to claim that it is not a fact about things themselves that everything must be an effect, but merely a fact about the nature of our experience.

‘individual’ is an object of thought, not an object (or even a possible object) of experience; it is an idea. It is impossible for us to experience any object as an individual. We can, however, if we are transcendental idealists, think of objects of experience as individuals without contradiction. In so doing we think of ourselves as members of an intelligible world. For we think of them as intelligible (rather than phenomenal) and we are related to them. This is what lies behind Kant’s claim in the *Groundwork* that:

By *thinking* itself into a world of understanding practical reason does not overstep its boundaries, but it would certainly do so if it wanted to *intuit* or *feel* itself into it. (4:458)⁷⁶

In addition, not only can we *think* ourselves into a world of understanding (an intelligible world), we can also choose to be members of such a world. For although we cannot *intuit* or *feel* ourselves into such a world, Kant believes that we can *will* our way into it. For although the *thought* of a world of understanding cannot be an object of my *faculty of intuition*, nor of my *faculty of feeling*⁷⁷, it can be an object of my *faculty of desire*, or will. This will be the topic of the following chapter.

⁷⁶ This claim should be read in conjunction with Kant’s claim in the preface of the second edition of the *Critique of Pure Reason* that, “even if we cannot **cognize** these same objects [objects of experience] as things in themselves, we at least must be able to **think** them as things in themselves” (Bxxvi).

⁷⁷ Kant perhaps changed his mind about this question.

Chapter Three

The Kingdom of Ends as a Community of Spirits:

Kant and Swedenborg

Introduction

I have argued that, for the mature Kant, the idea of a kingdom of ends is an idea of pure reason, being the idea of an intelligible world, or community. In this chapter I will examine the genesis of Kant's conception of a kingdom of ends. I will argue that Kant first started to think of morality in terms of striving to be a member of a kingdom of ends, understood as an ideal community, in the early 1760s, and that he was influenced in this by his encounter with the Swedish mystic Emanuel Swedenborg. Swedenborg wrote volumes about his visions of heaven and hell, and in 1766 Kant published a book on Swedenborg, *Dreams of a Spirit Seer*, a commentary on Swedenborg's magnum opus, *Heavenly Secrets*. Most commentators take Kant's attitude towards Swedenborg to have been entirely negative, and argue that, at the most, Kant's encounter with him had a purely negative impact upon his development, inducing him to reject certain of his early metaphysical positions. In this chapter, I will argue that Swedenborg had a positive influence upon Kant's development, particularly upon his ethics, for Kant's conception of

a kingdom of ends is modeled upon Swedenborg's conception of heaven as a community of spirits governed by moral laws.

Kant's engagement with Swedenborg in the early 1760s convinced him that it is possible for us to conceive of interaction between spirits. Up until this point he had believed that interaction was only possible between embodied beings. Swedenborg's descriptions of heaven as a community of spirits governed by moral laws, standing in non-spatial relationships to one another, provided Kant with a way of conceiving of a community of intelligible individuals. We can trace back the mature Kant's idea of a kingdom of ends, or an intelligible world, to Swedenborg's description of heaven as a community of spirits. That Kant's idea of a kingdom of ends is modeled on a particular conception of heaven should not be surprising if we remember that Kant often refers to the idea of a kingdom of ends as "the kingdom of God". In this chapter I will attempt to justify these claims.

Although we can trace the genesis of Kant's idea of a kingdom of ends to Swedenborg's account of heaven, I will argue in the chapters that follow that by the 1780s his idea of such a kingdom had departed from Swedenborg's conception of heaven in two major respects. Firstly although Swedenborg conceives of heaven as a kingdom governed by divine laws, he does not suggest that the members of the kingdom must be the "givers" of the laws. The mature Kant, in contrast, will argue that the idea of a kingdom of ends is the idea of a kingdom in which the members of the kingdom are the givers of the laws that provide the kingdom with its unity. That is, he believes that our idea of a kingdom of ends is the idea of a community of autonomous individuals. Secondly, Swedenborg

conceives of heaven as a community of spirits governed by laws of love, or what Kant will call laws of benevolence. The mature Kant, however, will maintain that we must conceive of the kingdom of ends as a political community, or ideal state, governed by juridical laws. Laws of benevolence, he will argue, are only possible in such a political community, and so we cannot conceive of a community governed solely by laws of love or benevolence.

In this chapter, then, I will argue that Swedenborg had a *positive* influence upon Kant's development. This is not to say that Kant was in any sense a follower or secret disciple of Swedenborg. Indeed he almost certainly believed that Swedenborg was deranged, and that his visions were probably due to some physiologically induced mental illness. However, although Kant believes that Swedenborg was almost certainly mad, this does not imply that he did not find his visions morally inspiring. My claim is that what Kant took from Swedenborg was the idea that morality demands that we develop a character that makes us a potential member of a kingdom of ends, or heaven considered as a community of spirits. Although Kant would later develop a more sophisticated account of the nature of such a community Kant's idea that morality involves striving to be a member of such an ideal community, and that the criterion for citizenship in such a community is the state of ones character, dates back to the mid-1760s and his engagement with Swedenborg.⁷⁸

⁷⁸ Kant was not the only person to be impressed with Swedenborg as a moralist. Coleridge, for example, writes that, "I can venture to assert that as a moralist, Swedenborg is above all praise" (quoted from Bellin & Ruhl 1985, p.ix).

In addition to the positive influence on Kant's ethical development, Swedenborg also had a positive influence on the development of Kant's theoretical philosophy, and in particular upon his belief in the ideality of space and time. For Swedenborg himself believed that the spatiality and temporality of objects of experience were due to our mode of perception and not due to the nature of the objects themselves, and Swedenborg believed that after death our doors of perception will be opened and we will experience things as they are in themselves. As we shall see, at least up until the early 1790s, Kant himself was committed to the position that after our bodily death we can hope for such a change in our form of intuition.

The claim that Swedenborg had a *positive* influence on Kant's development is definitely a minority position. For the majority of Kant scholars who attribute any influence to Swedenborg attribute a merely *negative* influence. The general structure of this *negative influence thesis* is that, priori to reading Swedenborg, Kant held a position that was similar to Swedenborg's. Upon reading Swedenborg, however, Kant realized the absurdity of his own earlier position; according to the *negative influence thesis*, then, Kant regarded Swedenborg's writing as a *reductio ad absurdum* of his earlier metaphysics. The two most significant recent proponents of the negative influence thesis are Laywine (1993) and Schönfeld (2000).⁷⁹ In this chapter, in addition to providing evidence of a positive influence I will also demonstrate the weaknesses of these two accounts of the negative influence thesis.

⁷⁹ And as Schönfeld himself acknowledges that he is following Laywine in attributing a merely negative influence to Swedenborg I will concentrate primarily upon refuting her presentation of the position. See Schönfeld (2000), p.244.

In the 1750s the young Kant believed that interaction was only possible between spatio-temporally embodied individuals. The reason for this commitment was his belief that interaction is only possible between impenetrable things (conceived of as centers of force), and he believed that only spatio-temporally embodied beings can be impenetrable. As a consequence the young Kant was implicitly committed to the position that real interaction between disembodied spirits is impossible.

By the 1780s, however, Kant has radically changed his position. For the mature, critical Kant maintains that real interaction is intelligible rather than phenomenal. He believes that we can only conceive of real interaction between intelligible beings, that is, between individuals conceived of as not subject to the spatio-temporal conditions of experience. In the language of the young Kant, then, the critical Kant maintains that real interaction is only possible between (disembodied) spirits.⁸⁰

Kant changed his position in the early to mid 1760s, and what provoked him to change his position was his engagement with the Swedish spiritualist Emanuel Swedenborg. Kant came across Swedenborg in the early 1760s and in 1766 published a book on his work, *Dreams of a Spirit-seer Elucidated by Dreams of Metaphysics*.⁸¹ Swedenborg was a mystic who wrote voluminously about his visions of the spiritual world. Kant clearly

⁸⁰ The mature Kant himself explicitly identifies the intelligible world with the ‘spiritual world’ in his metaphysics lectures from the early 1790s (over 10 years after having written the *Critique of Pure Reason*. See, *Metaphysik K2* where he identifies the intelligible world <*mundus intelligibilis*> with a spiritual world <*mundus pneumaticus*> (28:775). The idea of a ‘spirit’ is the idea of a disembodied (i.e. non spatio-temporal) individual, so it is the idea of an intelligible, rather than a phenomenal, being.

⁸¹ Which I shall refer to henceforth as *Dreams*.

thought that Swedenborg was deranged. There was, however, something valuable about his descriptions of his experience of the spiritual realm, for it suggested to Kant a way of conceptualizing intelligible interaction. For Swedenborg describes the spirit world as governed by spiritual laws with spirits as the locus of spiritual (or moral) forces, excluding or attracting one another on the basis of the state of their moral characters. Although spirits do not exist in space/time, they do stand in relations to one another, and there is something analogous to space in the spirit world for there is a 'moral distance' between spirits, which depends upon the respective states of their characters. In reflecting upon Swedenborg's account of the spirit world, Kant discovered a means of conceiving of spirits (or intelligible individuals) as impenetrable and standing in relations to one another, without having to think of them as embodied or necessarily spatio-temporal. This was an essential step in Kant's development for it provided him with a way of conceptualizing his moral ideal: a kingdom of ends as an intelligible world of individuals in interaction. Further reflection also led him to the conclusion that the only way of conceiving of such an intelligible world is as a community of autonomous agents. Before discussing Kant's engagement with Swedenborg, I will begin by justifying my claim that Kant changed his position between the 1750s and the 1780s.

(3a) Kant's Change of Position

The young Kant conceived of individuals as centers of forces and as a result believed that individuals must be impenetrable. In addition he concluded that this meant that all individuals, if they are to interact, must be spatio-temporally embodied.

In this I agree with Laywine (1993), who argues that as early the *True Estimation of Living Forces* (of 1747),

Kant claims, in effect, that the soul occupies a place not primarily because it is embodied, but because it can produce change of state in things other than itself. In short, the soul has a place by reason of its outwardly directed activity. . . This is his view not only in the *True Estimation*, but also in the *Nova Dilucidatio* [*New Elucidation*]. (p.45)

Discussing the *Physical Monadology* (of 1756), a work which Laywine believes expresses the same basic position of the *New Elucidation* of the previous year, Laywine explains that in this work Kant's position is that,

An element fills space by resisting every effort of every other element to penetrate the sphere of its activity. Thus elements fill space by making themselves impenetrable to one another. Unless we can show that the force whereby a soul is present in space is different from an element's force of repulsion, Kant is faced with the odd conclusion that the soul [or spirit] is impenetrable. (p.49 – my addition in square brackets)

The strongest textual evidence for the claim that the young Kant believed that only embodied individuals can really interact is to be found in the *New Elucidation*. In this work Kant attacks the doctrine of pre-established harmony, arguing that that if individual substances really were isolated worlds unto themselves, then it would be impossible for them to undergo any alterations of state. Given the fact that individuals do undergo alterations, then, they must really interact. He continues by noting that,

Our demonstrations [that change is impossible if we accept pre-established harmony] furnishes the opinion that some kind of organic body , must be attributed to all spirits whatever with powerful evidence of its certainty. (1:412)⁸²

⁸² In this passage Kant seems to be advocating a position he imputes to Leibniz, namely that every soul (monad) must have a material vehicle. This is a position the mature Kant clearly rejects. Thus is the early

This suggests that at this point Kant believed that embodiment was necessary for interaction, for the argument Kant is alluding to seems to be something like the following: (1) An individual substance can only undergo a change of states if it really interacts with other substances. (2) Spirits change their states. (3) Only embodied substances (that is, substances “to which some organic body can be attributed”) can really interact. Therefore, (4) spirits must be embodied. The conclusion Kant draws makes no sense unless he is implicitly assuming something like premise three.

The young Kant, then, seems to have believed that real interaction is impossible between disembodied spirits. By the time of the *Critique of Pure Reason*, however, he has clearly changed his mind. Thus in his metaphysics lectures from 1782-3, lectures given between the publication of the first and second editions of the *Critique*, Kant can claim that,

The world must also have only one cause. The connection <nexus> of substances is on that account to be thought possible only as derivative, but with that not as ideal, but rather concurrently as real. This proof holds, however, only for the noumenal world <mundus noumenon>. In the phenomenal world <mundus phaenomenon> we do not need it, for it is nothing in itself. Here everything is interaction <commercio> in virtue of space. The systems of occasional and predetermined harmony take place only in the sensible world. (29:868 *Metaphysik Mrongovius*)

Here Kant argues that real interaction occurs only in the intelligible world, and that there is no real interaction in the phenomenal world. In the language of the young Kant, this would be to claim that real interaction is only possible between disembodied spirits. The

1780s he will claim that, “the opinion of Leibniz, that the soul has here already and also will in the future a vehicle <vehiculum> of matter which is indestructible, is sensible and explains nothing” (*Metaphysik Mrongovius*, 29:920).

position that there is real interaction in the intelligible world is a consequence of Kant's claim that our idea of the intelligible world is the idea a community of individuals.

The young Kant conceived of 'force' and 'resistance' and 'impenetrability' as sensible concepts, applicable only to spatio-temporal beings. He believed that for two beings to resist one another implies that they must be in a spatial relationship to one another. Kant, however, did not remain committed to this view throughout his career. If he had remained committed to this position, he would have had to maintain that that spirits, or intelligible individuals could not really interact, for the mature Kant remains committed to the view that real interaction is only possible between beings that resist one another. The mature Kant, however, believes that resistance is a pure concept, being what he calls a predicable of the category of community.⁸³ As such it can be thought independently of the spatio-temporal conditions of experience. The same can be said of the concept of force which, Kant argues, is a predicable of the category of causality. Resistance and force, then, are pure concepts which can, of course, be applied to objects of experience but which can be thought without reference to the (spatio-temporal) conditions of experience. The fact that resistance and force are pure (unschematized) concepts implies that we can *think* of individuals resisting each other without having to think of them as spatially embodied. The germs of this view can be traced back to Kant's reading of Swedenborg in the early 1760s. That is not to say that the view was worked out in any detail at this time. For Kant would only develop the table of categories in the late 1770s. Kant's reading of Swedenborg, however, stimulated him to think about the possibility of 'moral' or 'intelligible' forces and relations.

⁸³ See A82/B108.

Engaging with Swedenborg helped Kant to develop an account of the possibility of real interaction between spirits, an account that did not commit him to the (for him untenable) position that spirits are necessarily embodied. This account was, I believe, only worked out fully with his introduction of the table of judgments and categories in his critical period, which allowed him to explain how we conceptualize real interaction in terms of the disjunctive form of judgment. The mature Kant, then, could argue that real interaction does not require that individuals are (spatially) impenetrable, but merely that they resist one another (by means of ‘intelligible’ forces). For we are able to *think* of a multitude of non-spatio-temporal individuals resisting one another because resistance can be conceived of independently of any spatio-temporal conditions. Both ‘resistance’ and ‘force’ are, Kant believes, pure concepts, which, although they can be applied to objects of experience can be thought of independently of the spatio-temporal conditions of experience. As a result of this, Kant believes that we can conceive of individuals as exercising forces and resisting one another without thinking of them as (spatially) impenetrable. The realization, then, that Kant came to while reading Swedenborg was that it is possible to distinguish between physical forces and intelligible (what Kant refers to as ‘moral’) forces. This distinction helped him to see that it is possible to conceive of a spirit/intelligible world in which there is real interaction.

(3b) 1763-6 - Kant’s Engagement with Swedenborg

Kant read Swedenborg in the early 1760s and in 1766 published *Dreams of a Spirit-seer elucidated by Dreams of Metaphysics*, a book dealing with Swedenborg's eight volume *Arcana Coelestia*. In his *Arcana Coelestia* Swedenborg, amongst other things, recounts his visions of heaven and his experiences with the world of spirits. Kant's response to Swedenborg has puzzled many commentators. In particular, there is little agreement on Kant's ultimate attitude towards Swedenborg. I will argue that although Kant believed that Swedenborg was mad, he was deeply affected by this encounter. In particular I believe, following Kant's most recent biographer Manfred Kuehn, that Kant underwent a moral conversion during the period he was engaging with Swedenborg and I will argue that there is a relationship between his reading of Swedenborg and the nature of his conversion. In particular Kant was very congenial towards Swedenborg's 'modern' conception of heaven as a spiritual community and the idea that the spiritual [or intelligible] world is not somewhere we are transported after death, but is an intelligible community of which we are already members, although without being able to intuit it.

In addition Kant was also deeply struck by Swedenborg's suggestion that it is up to us to determine which type of spiritual community we belong to and that in choosing a particular (moral) character we are choosing to be members of a community of similar characters. This is reflected in Kant's account of how we go about making moral judgments. For, according to Kant, when we are thinking morally about what sort of character (maxims) we should adopt, we think about whether it would be possible to be a member of a community of individuals with such characters.

(3c) Swedenborg

Swedenborg was born in 1688, and was an important figure in enlightened Swedish intellectual life in the early 18th century. He died in 1772. Amongst other things he was a respected engineer, mathematician and scientist. He wrote important works on metallurgy, chemistry, mineralogy and astronomy, and published the first work in Swedish on algebra, as well as co-founding Sweden's first scientific journal, *Deadalus Hyperboreans*.⁸⁴ He also wrote a four volume scientific treatise on the brain, based upon his own anatomical studies in which he discovered the functions of the cerebellum, the pituitary gland and spinal fluid. In 1716 he was offered, but turned down, the professorship in mathematics at the University of Uppsala and instead accepted the position of Assessor Extraordinary to the Swedish Board of Mines, an important position he held for almost 30 years. All in all Swedenborg could be regarded as a typical man of the enlightenment. In 1736, however, he started to have mystical visions and eight years later, on the night after Easter, April 6-7, 1744, he had a major mystical experience, believing he had personally encountered God, face to face, who had opened up his soul and revealed the world of spirits to him and commissioned him to spread the word about the true nature of the spirit world. Concerning this experience, he writes that, "from that day I gave up the study of worldly science, and I labored in spiritual things. . . The Lord opened my eyes. . . so that in the middle of the day I could see into the other world, and in a state of perfect wakefulness converse with angels and spirits".⁸⁵ After this he gave up

⁸⁴ This, and the following, information about Swedenborg's life is taken from Benz (2002), Laywine (1993), McDannell & Lang (1988), Bellin & Ruhl (1985) and Horn (1997).

⁸⁵ Quoted from Bellin & Ruhl (1985), p.43.

his official position and concentrated on his spiritual writings. From this period onwards he had frequent visions of both heaven and hell, and wrote many books about his experiences.

After his death in 1772 his followers founded a Swedenborgian church, the Church of the New Jerusalem, which exists to this day. Perhaps the most famous immediate follower of Swedenborg was the English poet William Blake who, for a short time, was an active member of this church. Many Swedenborgian elements and references can be found in his poetry and his *Marriage of Heaven and Hell* is a (critical) response to Swedenborg's *Heaven and Hell*. Many early abolitionists were followers of Swedenborg.⁸⁶ He had an influence on the German Romantics, especially upon Goethe, Schelling and Novalis, and had a strong influence on both American popular and high culture in the late 19th and early 20th centuries. His views were popularized through popular works, such as Elizabeth Stuart Phelps huge bestseller *The Gates Ajar* (1868), and Helen Keller's *Autobiography*, and parodied by writer such as Mark Twain in his *Extract from Captain Stormfield's Visit to Heaven*, and he had a strong influence upon Emerson and the Transcendentalists. Even the sober William James is known to have carefully read many of his books. At the very least, then, Swedenborg should be regarded as an interesting, if marginal, figure in our cultural history.

In addition to recounting his experiences with spirits, Swedenborg wrote volumes of inspired biblical interpretation. He believed that the bible has both an external and an

⁸⁶ One reason for this was Swedenborg's belief that Africans led a purer more spiritual life than Europeans, and that in the afterlife they were to be found in the highest heavens.

internal sense that he had been granted divine insight into this true internal sense of the divine word, and believed that his true vocation was to spread this inner word.

To understand Swedenborg's practice of biblical exegesis, and the importance he placed upon it, we must understand something about his doctrine of 'correspondences' and his account of the ages of mankind. At the heart of Swedenborg's theology lies his doctrine of 'correspondences'. According to this doctrine everything we experience (spatio-temporally) in this life 'corresponds' to something in heaven, which for Swedenborg is understood to be an organic community of angels. The most frequent metaphor Swedenborg offers to explain this doctrine is in terms of the human face. When we look at someone's face we can see their joy or sadness. Their outer appearance reveals their inner emotional state. The phenomenal world has the same relationship to the spiritual world as the expression on a human being's face has to their inner emotional state. Swedenborg believes, then, that the phenomenal world is, in effect, the face of heaven. Unfortunately, he believes that in our current fallen state we are not able to see it in these terms. Thus Swedenborg explains that,

We can see in the human face what correspondence is like. In a face that has not been taught to dissimulate, all the affections of the mind manifest themselves visibly in a natural form, as though in their very imprint, which is why we refer to the face as "the index of the mind." This is our spiritual world within our natural world. (Heaven & Hell, #91)

Although we are unable to immediately experience the natural world as the face of the spiritual world, there was a time when human beings could. To understand the importance Swedenborg places on his inspired biblical interpretation, it is necessary to

understand his account of the gradual fall of mankind. His simplest account of this falling away of mankind from heaven is to be found in *Heaven & Hell*, and this account of the stages of this fall is based upon Ovid's account in *Metamorphoses* of the three ages of mankind.

Swedenborg maintains that the earliest human beings were "heavenly people" who could read the heavenly significance of phenomenal events and objects in the same way that we can read a face. Thus he explains that the first age of mankind was the "Golden Age" and that at this time humans,

Thought on the basis of actual correspondences, and . . . the natural phenomena of the world that greeted their eyes served them as means for thinking in this way. Because they were of this character, they were in the company of angels and talked with them. (*Heaven & Hell*, #115)

In the Golden Age, which for Swedenborg was the age of Adam, humankind was face to face with heaven, or the community of angels.⁸⁷ After the fall, however, humankind became more separated from heaven and gradually lost this "face to face" connection with the heavenly angels. In the following age, which Swedenborg calls the Silver Age, mankind had not lost all connection to heaven. In this age,

People did not think from actual correspondences but from a knowledge about correspondences. There was still a union of heaven with humanity, but not such an intimate one.

After the fall, then, humans lost the ability to intuit heaven, but they retained an ability to understand the relationship between the phenomenal and the heavenly. In the age of the

⁸⁷ Swedenborg reads the bible symbolically, and believes that 'Adam' does not refer to a particular individual, but to an age of mankind.

old testament prophets mankind had lost the ability to intuit the phenomenal world as the face of heaven, but they still had knowledge of these correspondences, and this knowledge was collected in the old testament. The bible, then, explains these correspondences.

In the following age, the Bronze Age, this knowledge was replaced with a mere *familiarity*. Thus Swedenborg explains that in this age came people who “were indeed familiar with correspondences but [who] did not do their thinking on the basis of their knowledge of correspondences” (*ibid.*). This familiarity consisted in the ability to understand the true spiritual meaning of the bible.

In our age, however, even this familiarity has been lost, for

Humanity became more and more externally minded and at last physically minded. Then the knowledge of correspondences was completely lost, and with it any awareness of heaven and of its riches. (*ibid.*)

Swedenborg’s mission in life he believes is, at the very least, to restore our familiarity with heaven and its riches, for he was granted an intuition of the heavenly in order to be able to interpret the true spiritual meaning of the bible, and his magnum opus, Heavenly Secrets, the eight volumes of which Kant read and responded to, is an attempt to do just this.

For Swedenborg, then, the bible is like a textbook on physiognomy, but a textbook we do not know how to read. In the Bronze Age, however, people could understand it and use it as such. They were in a position similar to that of an imagined alien visitor to this planet,

a visitor who understands and feels human emotions, but is unable to see from looking at peoples faces how they are feeling. The bible is like a manual that can be referred to make judgments about what emotional states certain facial expressions signify. An alien visitor who met someone who was smiling, and, having checked the manual, could make the judgment that the person was happy. He would not see the person's happiness, but could make a judgment about it. The ancient readers of the bible were in a similar position. Unlike Adam, they could not see the heavenly in the phenomenal, but they could, by using the bible, obtain knowledge of, or at least familiarity with, the heavenly. Gradually, however, humankind became even more separated from heaven, and in the modern world we cannot even understand the true inner meaning of the bible. Swedenborg, however, believes that his eyes were opened to the true inner, spiritual meaning of the bible by God and he was assigned the task of acquainting the rest of humanity with this meaning.

As a result, much of his writing consists of bizarre symbolic biblical interpretation. An example, will give the reader some idea of his principles of interpretation. I will quote at length to give the reader some idea of Swedenborg's prose style.⁸⁸

Genesis 2:19-20 reads as follows: *“And Jehovah God formed out of the ground every beast of the field, and every fowl of the heavens, and brought it to the man to see what he would call it; and whatsoever the man called every living soul, that was the name thereof. And the man gave names to every beast, and to the fowl of the heavens, and to every wild*

⁸⁸ Which even Kant found to be “dull”: “The style of the author is dull” (2:360).

animal of the field; but for the man there was not found a help as with him.”

Swedenborg begins his commentary on this passage in the following terms,

By “beasts” are signified celestial affections, and by “fowls of the heavens” spiritual affections; that is to say, by “beasts” are signified things of the will, and by “fowls” things of the understanding. To “bring them to the man to see what he would call them” is to enable him to know their quality, and his “giving them names” signifies that he knew it. But notwithstanding that he knew the quality of the affections of good of the knowledge of truth that were given to him by the Lord, still he inclined to his own, which is expressed in the same terms as before – that “there was not found a help as with him.

That by “beasts” and “animals” were anciently signified affections and the like things in man, may appear strange at the present day; but as the men of those times were in a celestial idea, and as such things are represented in the world of spirits by animals, and in fact by such animals as they are like, therefore when they spoke in that way they meant nothing else. Nor is anything else meant in the Word in those places where beasts are mentioned either generally or specifically. The whole prophetic Word is full of such things, and therefore one who does not know what each beast specifically signifies, cannot possibly understand what the Word contains in an internal sense. But, as before observed, beasts are of two kinds – evil or noxious beast, and good or harmless ones – and by the good beasts are signifies good affections, as for instance by sheep, lambs, and doves. (p.76-7)⁸⁹

Each beast mentioned in the bible, then, signifies something specific. And so does every plant, element, name and number. Stone refers to faith or solid truths; water also refers to truth but “not in respect to its solidity, but in respect to its originality. . . and also to its reviving and cleansing properties.. . Birds refer to thoughts, and waterfowl to thoughts flowing like pure scientific truth” etc.⁹⁰ Swedenborg is particularly concerned with the importance of the inner meaning of numbers, arguing that, “it is clearly evident that

⁸⁹ Kant jokingly compares Swedenborg’s inspired method of interpretation to the play of the imagination which is at work in those who “discover the Holy Family in the irregular patterns of marble, or monks, baptismal fonts and organs in stalactites and stalagmites, or even the discovery by the mocking Liscow on a frozen window-plane of the triple crown and the number of the beast – none of them things which anyone else would see unless their heads were already filled with them beforehand” (2:360).

⁹⁰ These examples are from Solovyov (1997), p.4.

whatever numbers are used in the Word never mean numbers” (p.370). And, of course, Swedenborg has been granted special insight into these hidden meanings.⁹¹

Kant clearly thought Swedenborg was mad, and in *Dreams*, he declares that he would not blame the reader for regarding spirit-seers such as Swedenborg as “candidates for the asylum” (2:348).⁹² Many readers have taken Kant's attitude towards Swedenborg in *Dreams* to be entirely negative. However, although he was a sworn enemy of inspired interpretation, and was skeptical of any appeal to revelation and special insight, his attitude towards Swedenborg's visions is ambivalent, for his general attitude towards stories of the supernatural is not one of dogmatic rejection, but a skeptical agnosticism.⁹³ Thus he concludes the first part of *Dreams* with an assertion of his ignorance, which, he claims,

[P]revents my venturing wholly to deny all truth to the many different ghost-stories which are recounted, albeit with a reservation which is at once commonplace but also strange: I am skeptical about each one of them individually, but I ascribe some credence to all of them taken together. (2:351)⁹⁴

⁹¹ To us such views may seem ridiculous, and may be the source of an amused chuckle. In Kant's day, however, such views were far more mainstream. For the idea that biblical interpretation required special insight, provided by divine grace, was a standard feature of 18th century pietist doctrine. In understanding Kant's attitude to Swedenborg we should keep this fact in mind. Kant, of course, was a champion of the enlightenment, and so was a sworn enemy of such enthusiastic doctrines. But they would have appeared to him as far less abnormal than they do to a 21st century educated reader.

⁹² And he jokingly suggests that Swedenborg's visions may have been the result of misdirected wind, quoting Hudibras' opinion that: “if a hypochondriacal wind should rage in the guts, what matters is the direction it takes: if downwards, then the result is a f---; if upwards, an apparition or an heavenly inspiration” (2:348).

⁹³ Contemporary readers of Kant were not so quick to judge Kant's attitude as entirely negative. Thus Mendelssohn (1767), in his review of *Dreams*, writes that Kant's book, “occasionally leaves the reader in doubt about whether Mr. Kant wished to ridicule metaphysics or whether he intended to praise clairvoyance” (quoted from, Schönfeld 2000, p.181). And Oetinger, the founder of Swabian theosophy, wrote to Swedenborg on December 4, 1766, that, “we have a book, “Dreams of a Spirit-Seer,” that is full of lofty praise, but at the same time, in order not to seem fanatical [*schwärmerisch*] is equally full of derogatory remarks against you” (quoted from, Dole 1997, p.3).

⁹⁴ It appears that Kant is speaking *in propria persona* here, for in a letter to Moses Mendelssohn, written in 1766, after the publication of *Dreams*, he claims that, “It was in fact difficult for me to devise the right style

I suggest that although Kant had no time for Swedenborg's inspired interpretation, and was deeply unsympathetic to his doctrine of correspondences, he was profoundly affected by the content of Swedenborg's visions and that regardless of Kant's appraisal of Swedenborg's mental state, Kant's engagement with him had a profound effect upon Kant's development. For, following Schneewind and Kuehn, I believe that Kant developed the essentials of his mature ethics around 1764-5, while he was engaged with Swedenborg, and the fact that he arrived at this position at precisely the time he was engaging with Swedenborg is not coincidental, for he was drawn to Swedenborg's 'modern' conception of heaven as a society or community of spirits. And in Swedenborg's vision of heaven as a community of angels we find the genesis of Kant's idea of a kingdom of ends as an ideal community that we should strive to be members of.

(3d) Kant's Encounter with Swedenborg: The Facts

What do we know of Kant's engagement with Swedenborg? At the very least we know that Kant was seriously interested in Swedenborg between 1763 and 1766.⁹⁵ Establishing

with which to clothe my thoughts, so as not to expose myself to derision. It seemed to me wisest to forestall other people's mockery by first of all mocking myself; and this procedure was actually quite honest, since my mind is really in a state of conflict on this matter. As regards the spirit reports, I cannot help but be charmed by stories of this kind, and **I cannot rid myself of the suspicion that there is some truth to their validity. . .**" (10:70 – my emphasis).

⁹⁵ In addition Kant had positive things to say about Swedenborg during his metaphysics lectures ten years later in the mid 1770s. See 28:288-9. He also refers positively to Swedenborg in his lectures of 1792-3, see 28:690.

these dates is important, for they coincide with what one commentator has described as Kant's moral conversion of 1764.⁹⁶

Kant's first known reference to Swedenborg is found in a letter to Charlotte von Knobloch probably written in 1763. Kant begins the letter by explaining his attitude towards the paranormal, claiming that no one is in a position to accuse him of having a "mystical bent" or of having a "weakness for giving in easily to credulity" (10:43). And, although he does not "see such things as impossible", he used to be inclined to regard ghost stories and tales about spirits with skepticism. He continues, however, with the claim that: "That was my position for a long time, until I became acquainted with the stories about Herr Swedenborg" (10:44). Kant's skepticism about the paranormal, then, has been shaken by the stories about Swedenborg that have been relayed to him.

Intrigued by these stories he attempted to start a correspondence with Swedenborg and induced a number of his merchant friends to speak with him.⁹⁷ After explaining this (mediated) interaction, Kant continues his letter by recounting a number of the stories he has heard about Swedenborg. The incident that seemed to Kant "to have the greatest weight of any of these stories and really removes any conceivable doubts" (10:46) concerns a fire in Stockholm.⁹⁸ This fire occurred in 1756, while Swedenborg was in Gothenburg, about fifty miles from Stockholm. Swedenborg was at a party with about 15

⁹⁶ See Kuehn (2001), p.171. I will return to the question of Kant's moral conversion in the following section.

⁹⁷ And reports that, "He [Swedenborg] told my friend without any reservation that God had given him a wonderful power enabling him to communicate with souls of the dead whenever he pleased" (10:45).

⁹⁸ This story is also recounted in *Dreams* (2:355-6). Although, in this published work Kant is more skeptical about the veracity of the story.

other people. At about 6pm he started to look worried and explained to the other guests that he had had a vision that a fire had just started in Stockholm and was spreading fast, and he was worried that it would burn down his own house. Two hours later, however, he announced with relief that the fire had been put out, but had reached within three doors of his house. The story of Swedenborg's vision spread through Gothenburg's polite society that evening and even reached the Governor, who called him to his mansion and questioned him about the details of his vision. Swedenborg's vision occurred on Saturday night. On Monday evening a letter arrived from the merchants' guild in Stockholm describing the fire in the exact same terms as Swedenborg had. Kant concludes his account of this story by asking, "What objections can one raise against the authenticity of such a story?", for,

The friend who wrote me this investigated the whole matter personally, not only in Stockholm but as recently as two months ago in Gothenburg. He is very well acquainted with the most distinguished families in Gothenburg where everyone concerned told him the same story about this incident and most of the eyewitnesses of 1756, which is not so long ago, are still alive today. (10:47)

Here then was a supernatural story attested to by reliable sources. And, intrigued by this, Kant finishes his letter by informing von Knobloch that he, "eagerly awaits the book Swedenborg intends to publish in London. All arrangements have been made so that I will receive it as soon as it leaves the press" (10:48).

On November 6, 1764, (probably about a year after Kant's letter to von Knobloch), Kant's friend Hamann wrote to Mendelssohn that Kant, "was planning to review the

Opera Omnia of a certain Swedenborg [*sic*]”.⁹⁹ The work Kant had been reading was Swedenborg’s eight volume *Heavenly Secrets*, and his response, *Dreams of a Spirit Seer*, was published in 1766. Kant’s remarks on Swedenborg in this book are less flattering than in the letter to von Knobloch. He describes *Heavenly Secrets* as “eight quarto volumes stuffed full of nonsense” (2:360), and in his preface explains that *Dreams* was written because “the author went to the expense of purchasing a lengthy work, and what was worse, he put himself to the trouble of reading it, as well!” (2:318). What Kant found most tiresome in Swedenborg’s opus was his interminable biblical exegesis, and he writes in *Dreams* that, “none of these visionary interpretations are of any concern to me here” (2:360). The interspersed accounts of Swedenborg’s spiritual visions, however, were quite stimulating. Thus in *Dreams* he focuses exclusively on Swedenborg’s visions of the spirit world, explaining that,

It is only in the *audita et visa*, in other words, only what his own eyes are supposed to have seen and his own ears to have heard, which we are chiefly concerned to extract from the appendices attached to the chapters of his book. (2:360)

Although Kant's comments on Swedenborg in dreams are often negative, we shall see that he does have very positive things to say about Swedenborg in his later metaphysics lectures.

1764 – Kant’s moral “rebirth”

⁹⁹ Quoted from Kuehn (2001), p.171.

A number of important Kant scholars now believe that Kant had worked out the basis of his mature ethical position by the mid-1760s, and that this coincided with some sort of personal ‘moral conversion’. I believe that this story is basically correct. Following Lehman’s (1969) suggestion that Kant underwent a ‘life crisis’ in 1764,¹⁰⁰ Kuehn (2001), in his excellent new biography of Kant, argues that in 1764 Kant underwent a ‘moral conversion’. He writes that,

profound changes that took place in 1764. The elegant *Magister* with a somewhat irregular and unpredictable lifestyle changed into a man of principle with an exceedingly predictable way of life. He became like [his friend] Green. (p.156)

Schneewind (1998) also places a great emphasis on this period. Examining the development of Kant’s ethics, Schneewind provides a “story that now seems to make the best sense of the available evidence” (p.486). He argues that “the central point” of this story “is of course the claim that Kant had arrived at the essentials of his distinctive view of his morality by 1765” (*ibid.*). I agree with both Schneewind and Kuehn that 1764-5 marks an important turning point in the development of Kant’s ethics.

Kuehn suggests that Kant’s ‘moral conversion’ coincided with three important events in his life: (1) His 40th birthday on April 22nd, 1764,¹⁰¹ (2) the death of his best friend

¹⁰⁰ Lehman (1969), p.412.

¹⁰¹ Kuehn writes that, “On April 22, 1764, Kant turned forty. This was a significant event, at least in Kant’s own view of life. According to his psychological or anthropological theory, the fortieth year is of the greatest importance. . . [For] Kant believed that it is in our fortieth year that we finally acquire a character” (p. 144).

Johann Daniel Funk in April 1764, and (3) the development of his friendship with the English merchant Green.¹⁰²

I agree that the first two events probably played an important role in Kant's conversion. The development of Kant's friendship with Green, however, could not have played a role because, according to Kuehn, Kant did not meet Green until 1766, or at the earliest in 1765, after the date of his conversion. In addition, what seems to be missing from Kuehn's list is the fact that it was precisely at this time that Kant's interest in Swedenborg was at its peak.

Kuehn convincingly argues that at this time Kant was thinking deeply about the state of his character, and that his moral conversion involved a deep change in his character, or, to use Kant's own terminology, the conversion involved the *establishment* of a character. This focus on character (or what Kant calls one's 'disposition' or 'intelligible character' in his mature writings) lies at the heart of Kant's ethics for he believes that the choice of maxims is, in effect, a choice of character.¹⁰³ Thus, in his *Anthropology*, Kant explains that,

sometimes people say that a person has simply character (a moral character) which defines him as an individual and no one else. . . [such a moral character] is the distinguishing mark of a reasonable being endowed

¹⁰² Although Kuehn never explicitly makes the argument, the impression one gets upon reading Chapter four of Kuehn's illuminating biography of Kant is that Kant's moral conversion of 1764 was somehow influenced by his friendship with Green. Thus, for example, Kuehn remarks that in 1764, Kant "became more like Green" (p.156). However, as Kuehn himself points out elsewhere (p.154) Kant did not meet Green until 1766, or perhaps 1765, a year or two *after* his 'moral conversion'! This suggests that Kant's change in character was not somehow a result of this friendship, but, instead, that the change in character is what made his "deep moral friendship" with Green possible.

¹⁰³ Kuehn (2001), quite nicely emphasizes the importance of the idea of character in Kant's ethics by suggesting that maxims should be defined as "character-constituting principles" (p.147).

with freedom. The man of principles has character. Of him we know definitely what to expect. He does not act on the basis of his instinct, but on the basis of his will. (*Anthropology*, 7:285)¹⁰⁴

To have character, then, is to be a man of principles, and this is the distinguishing mark of a reasonable being who has a will rather than merely acting upon instinct. A little later Kant explains that his conception of character is to be distinguished from the usual understanding of the term, which “understands by character those qualities which accurately describe a person, be they good or bad” (7:292). For Kant, in contrast, moral character is not the sort of thing that can be good or bad; it is the sort of thing that one possesses or does not possess. Simply to have a character is “rare” and “admirable”, and he writes about the idea of character in the same language he uses to describe the idea of a good will in the *Groundwork*, writing, for example, that “character has an inner value and is above all price” (7:282) and that “having a character is the minimum requirement that can be expected of a rational person, and at the same time also the maximum of his inner value (of human dignity)” (7:295).

One is not born with character, but must, Kant believes, acquire it, and he writes that one can “take it for granted” that,

the establishment of character is, similar to a kind of rebirth, a certain solemn resolution which the person himself makes. This resolution and the moment at which the transformation took place remain unforgettable for him, like the beginning of a new epoch. This stability and persistence in principles can generally not be effected by education, examples, and instruction by degrees, but it can only be done by an explosion which suddenly occurs as a consequence of our disgust at the unsteady condition

¹⁰⁴ A few pages later he writes, in similar vein, that, “to have a character relates to that property of the will by which the subject has tied himself to certain practical principles which he has unalterably prescribed for himself by his own reason” (7:292).

of instinct. . . Wishing to become a better person in a fragmentary manner is a vain endeavor because one impression fades away while we labor on another. The establishment of a character, however, is absolute unity of the inner principle of conduct as such. (7:284-5)

Kant here writes as if he is speaking from experience. And we know that in the early 1760s he radically transformed his own lifestyle. He went from being an unpredictable young man to being the “man of principle”, the Kant of legends whose regularity was so famous that it was said that you could check the accuracy of your timepiece when you saw him start his afternoon walk. It is not, then, unreasonable to assume that Kant is basing these words on his own experience and that his “new epoch” began around the time of his fortieth birthday, in the early 1760s while he was reading Swedenborg. Further evidence for such a dating is provided by Kant’s remark that, “perhaps there will be only a few who have attempted this revolution before their thirtieth year, and fewer still who have firmly established it before their fortieth year” (*ibid.*).¹⁰⁵ Following Kuehn I suggest that this stress on the importance of ones fortieth year is probably based upon Kant's own personal experience.

This importance of a sudden moment of rebirth played a central role in the theory and practice of 18th century Prussian Pietism.¹⁰⁶ Kant himself received a pietist education at the *Collegium Fridericianum*, so it is not surprising that he was open to the idea of a sudden moment of moral conversion, for “the teaching staff in [pietist] institutions placed a higher priority on a reform of the will than on scholastic attainment” and “regarded a

¹⁰⁵ See Kuehn (2001), pp.145-8, for further evidence that Kant thought that one’s 40th year was a significant moment in life.

¹⁰⁶ 18th century pietism had a strong influence on the development of what has become American-style “born-again” Christianity.

conversion as the foundation of study. Students who had not yet experienced a ‘breakthrough’ were expected to exhibit a repentant attitude and demonstrate that they were preparing to be ‘born again’” (Gawthrop 1993, p.164). Francke, perhaps the most influential Prussian pietist in the early 18th century, revolutionized Prussian education, and the schools influenced by his teaching (including Kant’s) placed a huge emphasis upon “breaking the child’s natural will” (*ibid*, p.156) in the hope of provoking such a re-birth experience.

Kant’s attitude towards pietism is complicated. By the time of his education, pietism had been institutionalized in Prussia and was, in effect, the state religion, and Kant did not enjoy his early education. To get ahead in the Prussian state bureaucracy (which included educational institutions) it helped if you professed the faith, which involved being able to appeal to some personal moment of conversion or “breakthrough” (*Durchbruch*). This, of course, resulted in much hypocrisy with students, and, for that matter, anyone in an official state position, being rewarded if they could offer a story of personal conversion.¹⁰⁷ It is clear that Kant was disgusted by this hypocrisy and in his account of the establishment of character offered in the *Anthropology* he makes it clear that he does not believe that such a breakthrough can be achieved as a result of education. In arguing this he is strongly disagreeing with pietist practice.

¹⁰⁷ Fulbrook (1983) explains that in early eighteenth century Prussia, at the time Kant received his education, “the need for pietist testimonials to obtain positions in church and state led to superficial conversion and regeneration according to the routinised general stages of pietist experience. Pietism, conceived as a spontaneous religion of the heart had become rationalized and mechanical as the orthodoxy of the state” (p.170) The hypocrisy of many so-called pietists was a common criticism at the time. Thus Fulbrook quotes Semler (1781), a contemporary of Kant’s: “Now suddenly people were all supposed to become pious, or re-born; this alleged aim is impossible if one doesn’t count in all the hypocrisy and fanaticism. The true purpose was, to give oneself airs, without work or scholarship, and to get in with the Duke and Court.” (*ibid*. p.171)

In addition, Kant also found morally objectionable the pietist practice of treating the re-birth experience of others as models to follow. One of the dominant forms of pietist literature was the conversion narrative, and these narratives were used as models to be emulated. Thus Semler, a contemporary of Kant's, explains that, for the pietist, "the story of one's own experience and edification became the rule to follow exactly".¹⁰⁸ Kant objected to the practice of taking a phenomenal model as an ideal to emulate. Thus he argues in his ethics lectures that,

An example is when a general proposition of reason is exhibited *in concreto* in the given case. . . All cognitions of morality and religion [however] can be set forth apodictically, *a priori*, through reason. We perceive a priori the necessity of behaving so and not otherwise, so no examples are needed in matters of religion and morality. . . The examples must be judged by moral rules, not morality or religion by the examples. The archetype lies in the understanding. . . The reason why man would gladly imitate in matters of religion is that they fancy that if they behave as does the great majority among them, they will thereby constrain God, in that He cannot, after all, punish everybody. (27:333)

Imitating the behavior of others, then, is to undermine the purity of ethics. Rather than taking as our moral ideal the a priori ideal of being a citizen of a kingdom of ends we take as our ideal the empirical example of others. Given human weakness, taking the experience of another person, however, virtuous she may be, is to take something less than perfect as our model, and this makes it much easier for us to give excuses to ourselves. This is Kant's principled objection to the pietist practice of imitating the conversion experiences of others. Conversion, Kant believes, is something that we can experience personally, but it is not something to be imitated, for it is not something that

¹⁰⁸ Quoted from Fulbrook (1983), p.169.

we can choose. It is logically impossible to choose to be morally reborn, for we are morally reborn, Kant believes, when we choose to have a pure disposition. What it is to be reborn is to have chosen to have a pure disposition and in so far as we are attempting to choose to be reborn we are not choosing to have such a disposition.

Despite his reservations about pietist practice, it is clear that the pietist idea of a moral rebirth or “breakthrough” plays an important role in his ethics.¹⁰⁹ This is evident from the passage from the *Anthropology* already cited. The notion of a moral conversion is also a major theme in part 2 of Kant's *Religion*, and here Kant writes:

That a human being should become not merely *legally* good, but *morally* good (pleasing to God) i.e. virtuous according to the intelligible character (*virtus noumenon*) and thus in need of no other incentive to recognize a duty except the representation of duty itself, that, so long as the foundation of the maxims of the human being remains impure, cannot be effected through gradual *reform* but must rather be effected through a *revolution* in the disposition of the human being (a transformation to the maxim of holiness of disposition). And so a “**new man**” can come about only through a kind of **rebirth**, as it were a new creation and a change of heart. (6:47 – my emphasis)

Here Kant makes it clear that we cannot become moral gradually but that to become moral involves a sudden revolution and moment of rebirth. Kant himself hoped he was moral, and so must have believed that he himself went through such a revolution of character, and all the evidence points to the fact that this probably happened around 1764, at the time he was engaging with Swedenborg. Although the pietist notion of a

¹⁰⁹ In claiming this I am rejecting Kuehn's (2001) claim that, “It is absurd to claim that Pietism was a major influence on [Kant's] moral philosophy” (p.54). Of course, Kant was not a pietist. He found much of the actual, so called, pietist practice he saw around him distasteful and he strongly objects to the idea that our will must be broken so that we can subordinate ourselves to the will of God. However, Kant's belief that morality consists in the purity of our disposition is clearly influenced by the pietist ideal of purity of heart, as is his emphasis on the importance of moral conversion or rebirth.

“breakthrough” plays an important role in Kant's ethics, he secularizes this ideal. For the pietists this “breakthrough” involved subordinating one’s natural inclinations to the divine will, whereas, for Kant, it involves subordinating them to an idea, the idea of being a member of a kingdom of ends. In addition, repelled by the hypocrisy and “false pride” he saw around him, Kant believes that such a rebirth is a private matter, revealed to the world not through one’s words but through one’s actions. This disgust at the hypocrisy around him is, I suggest, one reason why the mature Kant, even though he believed himself to be morally reborn, felt disinclined to advertise the fact.¹¹⁰ Perhaps a deeper reason is that he believed that even if one has been morally reborn one cannot, or at least should not, present one’s own rebirth experience as a model to be emulated. Advertising his own moral rebirth might encourage others to attempt to emulate his rebirth experience, distracting them from the purity of the moral ideal within.

The death of his friend and his 40th birthday in 1764, then, left Kant thinking of death (and the possibility of an afterlife), the importance of friendship and the state of his own character. These events in his personal life left him very receptive to the ‘modern’ conception of heaven propounded by Swedenborg, with his conception of the afterlife as a community, one’s place in which is determined by the state of one’s character.

(3e) – Swedenborg’s Heaven and Kant’s Ideal of a Moral Community (Kingdom of Ends/Intelligible World)

¹¹⁰ For further information on the role of pietism in 18th century Prussia, see Gawthrop (1993), Fulbrook (1983) and Stoeffler (1973).

Introduction

If we believe that one's conception of the ideal state after death (if one has such a conception) reflects something deep about one's moral convictions, the fact that Kant found Swedenborg's conception of heaven appealing should, at the very least, tell us something about his ethical theory. In the case of Kant, I believe that the relationship between his 'image of heaven' and his ethics is particularly strong, for Kant believes that to be moral is to choose to be a member of an intelligible world and he is not adverse to identifying the idea of an intelligible world with the idea of "the kingdom of God", or the "kingdom of heaven".¹¹¹ One of the reasons for this is because Kant is drawn to the 'Swedenborgian' conception of heaven as a community and believes, with Swedenborg, that morally we should think of ourselves as already in heaven (or hell) but without realizing it, and we should believe that our spiritual location depends upon our choice of character.

In *Heaven a History*, McDannell & Lang (1988) present Swedenborg as a major manifestation of what they call the 'modern perspective on heaven'. Traditionally, they argue, the joy of blessed soul in heaven consisted primarily in the relationship of that soul towards God. According to the 'modern' conception, however, a major, if not the, joy in heaven consists in the interaction of the blessed. Thus, they argue that although,

The concept of a saintly community in heaven has a long tradition in Christian history, originating in the book of revelation. Christians acknowledged their belief in the "communion of Saints" each time they

¹¹¹ See, in particular, Part 3 of *Religion*. Also see the *Critique of Practical Reason* (5:137), where he explicitly makes such an identification, writing: "intelligible world (the kingdom of God)".

recited the Apostles' creed. However, what began during the Renaissance and more clearly in the seventeenth and eighteenth centuries was the recognition that heavenly happiness did not hinge on the vision of God but on the social interaction of the saints. No longer did the saints merely dance with the angels outside the celestial gates; they now enjoyed each other's company in the full sight of the divine. (p.211)¹¹²

McDannell & Lang base their analysis on both textual and iconographic sources. In the final sentence of this passage they are referring to the fact that in most medieval depictions (paintings and woodcuts, for example) of the last judgment the blessed may be depicted as paying an interest in, and interacting with, one another outside the gates of heaven. Once beyond the gates, however, they are nearly always depicted as focusing all of their attention upon the presence of the divine and not upon one another. Beginning with the renaissance, it is far more common to see the blessed depicted as interacting with each other, even in the presence of God. This trend towards depicting the state of the blessed as an idealized human community reached a peak, they argue, in the works of Swedenborg.

Johnson (1996) explains Swedenborg's conception of the spiritual world as follows:

The spiritual world consists of three realms: heaven, hell, and an intermediate realm that he calls the world of spirits. Heaven is populated

¹¹² McDannell & Lang identify four characteristics of the 'modern' conception of heaven, all of which they find in Swedenborg: "First, only a thin veil divides heaven from earth. For the righteous, heavenly life begins immediately after death. Concepts of purgatory or sleeping in the grave until the general resurrection are either denied or minimized. Secondly, rather than viewing heaven as the structural opposite of life on earth, it is seen as a continuation and fulfillment of material existence. . . . Thirdly, although heaven continues to be described as a place of "external rest", the saints are increasingly shown engaged in activities, experiencing spiritual progress, and joyfully occupying themselves in a dynamic, motion filled environment. The journey to God does not end with an admittance to heaven but continues eternally. Spiritual development is therefore endless. Finally, a focus on human love expressed in communal and familial concerns slowly replaces the primacy of divine love experienced in the beatific vision. Social relationships, including the love between man and woman, are seen as fundamental to heavenly life and not in conflict with divine purpose" (p183).

by angels and hell by demons, all of whom are the departed spirits of rational beings who formerly inhabited earth and other planets. The intermediate world of spirits is populated by both departed spirits and by the spirits of living, embodied beings. Every rational being holds a dual citizenship in both the material and the spiritual worlds. Each of us exists always-already in a relationship with a spiritual self, what we might call the “better angels” of our nature. This spiritual self is the soul, understood both as the animating principle of the body and as our moral personality. Since each of us already exists in the spiritual world, the departure of the soul to the spiritual world is not to be understood as a journey from one place to another. Rather, it is to be understood as a transformation of our mode of cognition from sensuous intuition, which shows us only the material world, to a spiritual form of cognition, which reveals to us the place we already occupy in the spiritual world. . . There are three main spiritual laws governing the spiritual world: divine love, divine wisdom, and “use”. . . Divine love is the most primordial pneumatic law. . . Each community in the spirit world consists of spirits who have developed similar “loves”, similar hierarchies of value, [and] similar moral characters or temperaments during their embodied existence. (p.4)

This depiction of heaven as an ideal human community struck a chord with Kant and, as we shall see, he advocates a very similar position in his metaphysics lectures. In addition, he was sympathetic to Swedenborg’s belief that it is up to us, and not God, to choose which spiritual community (either heaven or hell) we belong to through the choice of our character.¹¹³ Thus, Swedenborg (1995) writes that, “Heaven is in a man, and people who have heaven in themselves come into heaven” (p.319). Similarly, “the evil within a person is hell within him and after death, his greatest desire is to be where

¹¹³ As McDannell & Lang (1988) point out, “Swedenborg radically departed from the orthodox Christian belief in an individual and final judgment. The spirit, not God, ultimately decided where to spend eternity” (p.189).

his own evil is. . . Consequently the person himself, not the Lord, casts himself into hell” (1997, p.547).¹¹⁴

To conclude: Although it is tempting to dismiss Swedenborg as a lunatic from a bygone era, there is something decidedly ‘modern’ in his madness: *firstly*, in his conception of heaven as a community and, *secondly*, in his rejection of the idea of the last judgment as an external judgment, made by God, at or after our death. Kant was drawn to both of these views, both of which are incorporated into his mature ethics. In addition, from the theoretical perspective, his reflections on Swedenborg pushed him towards his critical distinction between the phenomenal and intelligible world.¹¹⁵

Kant and Swedenborg on the post-mortem condition – a “cleansing of the doors of perception”?¹¹⁶

¹¹⁴ The famous Russian philosopher Vladimir Solovyov, in his encyclopedia article on Swedenborg, recounts one of Swedenborg’s visions, that illustrates this position: “At this time my inner person was in the middle heaven. . . which consists of a community of spirits who love truth because it is good. In their presence I felt their strong influence on my heart and proceeding to it to my brain, and the thought occurred to me, Is there any way in which the Lord’s mercy could let devils remain in hell to eternity? Even while I was thinking about this, one of the angels of a just temperament flew down with uncommon speed to the throne region of the great Satan and at the Lord’s suggestion brought out one of the evil devils in order to grant him heavenly bliss. I was allowed to see, however, that as the angel rose into a heavenly sphere, the proud expression on his prisoner’s face changed into one of suffering and his body turned black. . . dreadful convulsions came over him. . . and he showed that he was suffering immense and unbearable pain. . . His misery touched me, and I begged the Lord to command the angel to let him go. When, with the Lord’s consent, he was released, he hurled himself down headfirst so impetuously that all I could see was how his extraordinary black heels flashed by. . . Then I was given the insight that anyone’s stay in heaven or hell depends not on the arbitrary will of God but on the inner state of one’s essential nature. . . In this way, I understood that the eternity of hell for people who arrive there for their own gratification is in complete accord with both the wisdom and the goodness of God” (1997, p.5).

¹¹⁵ Thus he concludes section one of *Dreams* with the observation that “from now on it will perhaps be possible, perhaps, *to have all sorts of opinions* about but no longer *knowledge* of such beings” (2:351). This claim is in line with his critical position that we can think of the intelligible world, but have no cognition/knowledge of it.

¹¹⁶ “If the doors of perception were cleansed everything would appear to man as it is” – William Blake, from *The Marriage of Heaven and Hell*. Blake was, for a time, a member of the Swedenborgian New

Swedenborg believes that, “Every man while living in the body is in some society of spirits and of angels, though entirely unaware of it.” (p.352). Kant holds a very similar view, believing that although we can only intuit ourselves as members of the phenomenal world we should think of ourselves as members of a spiritual or intelligible world. Thus, in the *Critique of Practical Reason* Kant argues the antinomies of pure reason (from the *Critique of Pure Reason*) are a labyrinth, to which transcendental idealism provides the key. In discovering this key, however, reason “further discovers what we did not seek and yet need, namely a view into a higher, immutable order of things **in which we already are**” (5:107 – my emphasis).

Further evidence that Kant conceived of the afterlife in these terms is to be found in his lectures on metaphysics. In these passages Kant not only claims that we should regard ourselves as now already members of a spiritual (or intelligible) community, but without being able to intuit it, but also that we should hope that upon our death our form of intuition will change and we will be able to intuit our membership. This view is clearly derived from Swedenborg, and Kant himself acknowledges this debt.

Kant's metaphysics lectures followed the structure of Baumgarten's metaphysics textbook, and Kant customarily discussed the question of death and the post-mortem condition at the end of his discussion of rational psychology. In the mid-1770s, before the publication of the *Critique of Pure Reason* in 1781, he could claim that,

Jerusalem church, and this poem was written as a response, and commentary on, Swedenborg's *Heaven And Hell*.

We have a cognition of the bodily world through sensible intuition insofar as it appears to us; our consciousness is bound to animal intuition; the present world is the interaction <*commercium*> of all objects, insofar as they are intuited through present sensible intuition. But **when the soul separates itself from the body, then it will not intuit the world as it appears, but rather as it is.** Accordingly **the separation of the soul from the body consists in the alteration of sensible intuition into spiritual intuition,** and that is the other world. The other world is accordingly not another location, but rather only another intuition. (*Metaphysik L1*, 28:296 – my emphasis)

Some commentators may think such views are pre-critical and are incompatible with his critical project. Kant, however, repeats this claim in his lectures throughout the 1780s and into the 1790s. Given the fact that many readers of Kant might be surprised by this commitment, it is worth quoting these passages at some length.

Thus in 1782-3, in a lecture course he gave between the publication of the first and second edition of the *Critique of Pure Reason*, Kant argues that:

Now we find ourselves already in the intelligible world, and each human being can count himself as belonging, according to the constitution of his manner of thinking, either to the society of the blessed or of the damned. He is now only not conscious of it, and after death he will become conscious of this society . . . We are now already conscious through reason of finding ourselves in an intelligible realm; after death we will intuit and cognize it and then we are in an entirely different world that, however, is altered only in form, namely, where we cognize things as they are in themselves. (*Metaphysik Mrongovius*, 29:919-20)

Here Kant once again suggests that we can hope for some form of intellectual intuition after death. The claim that we are “now already conscious *through reason* of finding ourselves in an intelligible world” should be understood as meaning that even though we are at present unable to *intuit* ourselves as members of an intelligible world we are able to

think of ourselves as members of such a world, for the idea of an intelligible world is an idea of pure reason. And Kant makes it clear, once again, that he believes that it is not irrational to hope that at some point we will have an intuition of our membership.

In his lectures in 1784 he repeats the claim that “the virtuous is already in heaven only he is not conscious of it” (28:445), and adds that,

Cutting off all further pondering on this is the best remedy, that we can say: another world means only another intuition of the same things, the sensible world thus entirely ceases for us . . . Now it is asked: will the soul exist as pure intelligence? But it is indeed that when it is not sensible. But one also cannot think how a being that is created should cognize things in themselves. We will thus presumably come only by degrees to a greater perfection of cognitions and have another kind of intuition in the same or in another world. Here no philosophy goes any further. (*Metaphysik Volkmann*, 28:446)

And in his lecture course from 1790-1 he once again repeats the claim that “the human being who is virtuous is in heaven, only he does not intuit it, but he can infer it through reason.” (28:593) He continues by adding that,

the transition from the sensible world into the other is merely the intuition of oneself. According to content it is always the same, but according to form it is different . . . One sees at once how limited is our knowledge of the state of the soul after death. This life shows nothing but appearances, another world means nothing other than another intuition, things in themselves are unknown to us here, but whether we will become acquainted with them in another world? We do not know. A pure spirit cannot exist merely as soul in the sensible world. As intelligence it does not appear in space, also not in time. (*Metaphysik L2* 28:593)

Thus, throughout the 1780s, the decade in which Kant wrote the *Critique of Pure Reason*, the *Groundwork* and the *Critique of Practical Reason*, Kant was committed to the claim that we can (and should) hope for intellectual intuition after death. Some commentators

may argue that we should not place too much emphasis on unpublished lecture notes jotted down by his students. However, there is much consistency in the notes and the doctrine Kant presents here is clearly not just Kant's summary of Baumgarten's position. And, in addition, there is also evidence in his published writings that Kant is committed to such a position. For example, in the preface to the second edition of the *Critique of Pure Reason* Kant talks of,

that remarkable predisposition of our nature, noticeable to every human being, never to be capable of being satisfied by what is temporal (since the temporal is always insufficient for the predispositions of our whole vocation) leading to the hope of a future life. (Bxxxiii)

Here Kant makes it clear that the future life we must hope for is atemporal, which, given Kant's account of time as a form of intuition, can only mean that he believes that we must hope for some change in our form of intuition (into a non-temporal form of intuition) after death.

This position is clearly derived from Swedenborg's claim in *Heavenly Secrets*, that,

Every man while living in the body is in some society of spirits and of angels, though entirely unaware of it. And if he were not conjoined with heaven and with the world of spirits through the society in which he is, he could not live a moment. . . The very societies in and with which men have been during the life of the body, are shown them when they come into the other life. And when, after the life of the body, they come into their society, they come into their veriest life which they had in the body, and from this life begin a new life; and so according to their life which they have lived in the body they either go down to hell, or are raised up into heaven. (p.352)

He was drawn to such a position because he believed that if we are to attempt to be moral we must have some hope that we can eventually have some awareness of our true moral

disposition (or, what he calls in the *Critique of Pure Reason* our ‘intelligible character’). For example, in the *Critique of Practical Reason*, Kant can argue that, “an upright man cannot be happy if he is not first conscious of his uprightness” (5:116). However to be upright is to have a moral disposition or intelligible character, and this is not the sort of thing that can be an object of our form of intuition. As a result we can have no knowledge of our uprightness. Thus Kant can write in the *Religion* that,

According to the law, each and every human being should furnish in his own self an example of [the] idea [of a human being morally pleasing to God]. And the required prototype always resides only in reason, since outer experience yields no example adequate to the idea; as outer, it does not disclose the inwardness of the disposition but only allows inference to it, though not with strict certainty. (Indeed, even a human being’s inner experience of himself does not allow him so to fathom the depth of his heart as to be able to attain, through self-observation, an entirely reliable cognition of the basis of the maxims which he professes, and of their purity and stability). (6:63)

Kant believes, then, that the virtuous man, if he is to be happy must have assurance of his uprightness. This, however, is impossible, given our form of intuition, because to be upright is to have a good intelligible character, and our intelligible character is not a possible object of (our form of) intuition. Kant also believes that we can hope to be happy. Therefore he concludes that we must hope that our form of intuition will change.

Interestingly, however, Kant does not continue to maintain that we must hope for a change in our form of intuition in his metaphysics lectures from the 1790s, and I suspect that he changed his position while writing the *Critique of Judgment*. A full examination of this issue would have to involve a careful interpretation of the *Critique of Judgment* and his short essay *The End of All Things*, published in 1794. My hypothesis is that in the

1790s he decided that in order to be assured of his uprightness, the virtuous man does not need to *intuit* his membership in the intelligible world, but could *feel* it. Such a *feeling*, as opposed to an intuition, of one's own uprightness (that is, a felt assurance of one's membership in a kingdom of ends) would be enough to make the virtuous man happy. We can be assured of our membership in such a world by experiencing the beauty of other (autonomous) individuals around us.¹¹⁷ We cannot hope to intuit their individuality and autonomy, but we can hope to feel it. For (a) the ideal of beauty is the (moral) human being¹¹⁸ and (b) the ideal aesthetic judge is the disinterested moral agent. I believe that in his account of the feeling of beauty in the *Critique of Judgment* Kant came to see a way of satisfying his hope for some awareness of our membership in the kingdom of heaven without having to appeal to the possibility of intellectual intuition after death. For, if we were perfectly moral (and hence perfectly disinterested) we would *feel* the beauty of those autonomous agents around us.¹¹⁹ A more detailed discussion of this interesting issue is beyond the scope of this dissertation.

¹¹⁷ Thus I believe that one of the major aims of the Critique of Judgment is to explain how there can be a visible expression of moral ideas. Thus, in his discussion of the ideal of beauty in section 17 of the *Critique of Aesthetic Judgment* Kant explains that “**the visible expression of moral ideas**, which inwardly govern human beings, can of course be drawn only from experience, but as it were to make visible in bodily manifestation. . . their combination with everything that our understanding connects with the morally good in the idea of the highest purposiveness – goodness of soul, or purity, or strength, or repose, etc. – this requires pure ideas of reason and great imagination. . .” (5:235 – my emphasis). The aim of the Critique of Judgment is to explain how moral ideas (such as the idea of an autonomous individual) can have a visible expression. That is, how it is possible to experience a visible body as an autonomous individual. In claiming this I agree with Guyer (1993) that “Kant did not look to moral theory to solve a problem in aesthetic theory; instead, he looked to aesthetics to solve what he had come to recognize as crucial problems for morality” (p.19). Although we disagree about exactly what moral problem Kant is attempting to solve.

¹¹⁸ “There is still a distinction between the normal idea of the beautiful and its ideal, which on the grounds already introduced can be expected only in the human figure. In the latter the ideal consists in the expression of the moral, without which the object would not please universally and moreover positively” (5:235).

¹¹⁹ Our experience of beauty, then, could, as Schiller suggests, “serve as a pledge in the sensible world of a morality as yet unseen” (Schiller 1967, p.15).

Swedenborg and the phenomenal/intelligible distinction

I believe that engaging with Swedenborg pushed Kant towards developing his phenomenal-intelligible distinction, and that Kant conceives of the intelligible world as a community of spirits in real interaction. Kant explicitly calls the intelligible world as the ‘spiritual world’. For example, in his metaphysics lectures from the early 1790s (over 10 years after having written the *Critique of Pure Reason*), he explicitly identifies the intelligible world (*mundus intelligibilis*) with the spiritual world (*mundus pneumaticus*) (*Metaphysik K2*, 28:775). Laywine also suggests that Kant’s reading of Swedenborg deeply affected him, and that this engagement led him to develop the phenomenal-intelligible distinction. However, the reasons she gives for this are very different from mine. And she attributes a very different conception of the intelligible world to the mature Kant than I do, for she believes that the mature Kant was committed to the position that there could be no real interaction in the intelligible world, for the idea of interaction between disembodied spirits is unintelligible.

Laywine (1993) maintains that Swedenborg, like the young Kant, also regarded spirits as necessarily embodied and spatio-temporal. According to Laywine, Swedenborg, in effect, functioned as a mirror to Kant. The young Kant was committed to the view that spirits interact, and as a result believed that they must resist one another and be impenetrable. As a result of this the young Kant concluded that spirits must necessarily be embodied. In reading Swedenborg, Laywine suggests, Kant recognized his own outlandish position reflected warts and all. And recognized that unless he clearly distinguished between the

phenomenal and the noumenal his position was equally outlandish. In the course of engaging with Swedenborg, then, Kant realized that he must clearly distinguish between the intelligible and the phenomenal world, and allow for real interaction only in the phenomenal world. And this was a position he maintained for the rest of his life.

Laywine maintains, then, that the critical Kant believed that real interaction is impossible in the intelligible world. Thus she argues that,

[K]ant apparently denies that pre-established harmony prevails in the Kingdom of Ends. But how can he deny this, given his remarks about the system of Leibniz in the Note to the Amphiboly in the first *Critique*? There he says that, if we use our pure concept to represent community in the intelligible world, we must apparently represent this world as one in which pre-established harmony prevails. So how might we possibly conceive of the Kingdom of Ends, if not as a system of pre-established harmony? (p.142)

Elsewhere, Laywine claims that in remark three of the *Amphiboly*, “Kant goes so far as to say that Leibniz was perfectly right to espouse pre-established harmony” (p.139), and that Kant’s position in the first *Critique* is that, “the metaphysician cannot reasonably conclude, even on his own dogmatic terms, that physical influx prevails in the intelligible world. Leibniz was right: the intelligible world presents a system of pre-established harmony” (p.140). This is clearly a misinterpretation of the position of the mature Kant, for there is ample textual evidence to show that the mature Kant conceives of the intelligible world as a community in which there is real interaction.¹²⁰ The only textual evidence she points to support her interpretation are Kant's comments on Leibniz in

¹²⁰ See chapter 1.

remark three of the *Amphiboly* of the *Critique of Pure Reason*, and she clearly misreads this passage.

In remark three of the amphiboly Kant explains the reason why Leibniz's "principle of the possible community of substances amongst themselves had to be predetermined harmony and could not be a physical influence" (A274/B330). Laywine claims that the reason Kant gives is that it is impossible for us to conceive of real interaction using the unschematized concept of community. This interpretation, however, is totally unwarranted. For she fails to read remark three in its historical context, and, in particular, in the context the disagreement between Wolff and Crusius about the relationship and distinction between the will and the understanding. Schneewind (1998) explains that at the heart of this disagreement is Crusius's belief that, "the Wolffians are. . . mistaken in thinking will and understanding to be only one power" (p.446).¹²¹ And Kant clearly sides with Crusius in this debate.

Thus, Kant continues remark three by explaining that the reason for Leibniz's commitment to pre-established harmony is because, for Leibniz, "everything is only internal, i.e. occupied with its own representations" (*ibid.*). And what he means by this is that Leibniz attributes to

substances no other inner state than the thought through which we internally determine our senses itself, namely **the state of representations**. This completes the monads, which are to constitute the fundamental matter of the entire universe, the active power of which, however consists merely in representations, through which they are

¹²¹ The following discussion of Crusius and Wolff is indebted to Schneewind.

properly efficacious merely within themselves. (A274/B330 – my underlining)

This, then, is the reason Kant points to as for why Leibniz was committed to pre-established harmony. And any educated contemporary reader would have understood what Kant was alluding to here, namely the fact that the Leibnizians reduced all faculties to a single power or faculty. The reason, then, why Kant believes Leibniz was committed to pre-established harmony was because he, like

Wolf [*sic*] wished to derive everything from the faculty of knowledge and defined pleasure and unpleasure as action of the faculty of knowledge. He also called the faculty of motivation a play of representations, and so merely a modification of the faculty of knowledge. (28:674)

Any contemporary educated German reader would have recognized that this is what Kant was alluding to in the *Amphiboly* when he claims that for Leibniz “the active power [of monads] consists merely in representations”. Anyone aware of this debate, then, should be able to recognize that in remark three of the *Amphiboly* Kant does not come anywhere near claiming that pre-established harmony must prevail in the intelligible world. All he is claiming is that if, like Leibniz and Wolff you assume that all human powers (or faculties) can be reduced to the power of representation *then* you must be committed to regarding individual as isolated (and thus to pre-established harmony). Kant, however, clearly rejects this assumption, for, following Crusius, he sharply distinguishes between our faculty of cognition and our faculty of desire.

Although Laywine’s account of Kant’s attitude towards Swedenborg in the 1760s is based upon a misinterpretation of his mature position we, can still learn something from

examining the problems with her account. In addition, her view have been influential on others working on Kant's development. For example, Schönfeld (2000) in his recent book *The Philosophy of the Young Kant*, accepts Laywine's interpretation without revision.¹²²

Thus, he writes that,

The inevitable consequence of the pre-critical project was that bodies and souls, or material and immaterial substances, are subject to the same laws. At the same time, the pre-critical project must not rule out the possibility of an afterlife – that is the possibility that material substances remove themselves from their physical embodiment and interact purely among themselves. . . What would such an immaterial community of souls look like? Because souls are substances that obey the same fundamental laws as bodies, the immaterial community of the souls must contain the same structure as the physical world. The *reductio ad absurdum* of the pre-critical project is Swedenborg's spirit-world – a world whose ghostly inhabitants are not even aware of their postmortal state because it looks and feels just like their old home¹²³. . . It is therefore correct to say (Laywine, 1993) that Kant found in the *Arcana Coelestia* a caricature of his own metaphysics. (p.244)

Schönfeld, following Laywine, believes that Kant regarded Swedenborg's work as the *reductio ad absurdum* of his own earlier position.

According to Laywine and Schönfeld, then, Kant found Swedenborg's writings to be ridiculous but also saw them as a mirror in which he could see reflected the absurdity of his own earlier position. This recognition provoked Kant to reflect upon his own earlier metaphysical commitments and to reject his earlier account of the spatiality of spirits and

¹²² And it should be noted that there are very few books on the development of Kant's views in the 1760s.

¹²³ This claim is also to be found in Laywine. What both Laywine and Schönfeld fail to recognize is that Swedenborg distinguishes between spirits and angels (and demons). He believes that even though immediately after death our form of intuition remains the same, and hence many spirits after death are not aware that they have died, over time ones form of intuition changes and one comes to recognize oneself as a member of either heaven or hell. Although neither Laywine nor Schönfeld notice this aspect of Swedenborg's theology, Kant himself, as we shall see, does. This makes it clear that Kant actually read Swedenborg's work quite carefully.

to carefully distinguish between the sensible and the intelligible in his next work – the *Inaugural Dissertation* of 1770. As Laywine explains it:

On Kant's own view, it would seem that the soul is an object of sensation in as much as we could collide with one. Now Swedenborg also represents immaterial things – angels and departed spirits – as objects of sensation. . . [On reading Swedenborg, Kant] was impressed by the general fact that he could not reasonably dismiss Swedenborg's reported conversations with angels and departed spirits so long as it was possible on his own view to collide with Spirits who had passed on to the hereafter. . . Kant did not find Swedenborg's work problematic just because it is all about angels and spirits. Kant himself was not troubled by admitting that it might be possible for such things to exist. Even in Dreams, he refuses to say [sic] that the existence of angels and spirits is impossible. . . The problem with Swedenborg was rather that the spirit-seer of Stockholm represents immaterial things as though they could be subject to the conditions of sensibility. (p.57)

Kant's response to this problem was, according to Laywine, to conclude that (a) spirits (or souls) cannot be subject to the conditions of sensibility, and as a consequence that (b) they cannot collide with one another and (c) that they cannot really interact. On my interpretation Kant drew almost the opposite conclusions, namely that, (a) the objects that we experience around us as subject to the conditions of sensibility can be thought of (although not intuited as) intelligible individuals (or spirits), (b) intelligible individuals can be thought of as centers of intelligible (moral) forces and as resisting one another, and, as a consequence of this, (c) intelligible individuals can be thought of as really interacting.

In the course of his engagement he found a way out of his dilemma, for Swedenborg's visions suggested to him that real interaction, although it involves resistance and forces, does not necessarily have to involve physical forces, which can only be applied to spatio-

temporal bodies. Indeed ten years after reading Swedenborg, Kant could still talk of Swedenborg's visions as 'sublime', and what he found so sublime about Swedenborg was that he clearly distinguished between the sensible world and the spiritual (intelligible) world. Thus, in his metaphysics lectures from the mid 1770s, ten years after his engaging with Swedenborg, Kant could argue that,

The thought of Swedenborg is in this quite sublime. He says the spiritual world constitutes a special real universe; this is the intelligible world <mundus intelligibilis> which must be distinguished from the sensible world <modo sensibilis>. He says all spiritual natures stand in connection with one another, only the community and connection of the spirits is not bound to the condition of bodies; there one spirit will not be far or near to the other, but rather there is a spiritual connection. Now as spirits our souls stand in this connection and community with one another, and indeed already here in this world, only we do not see ourselves in this community because we still have a sensible intuition; but although we do not see ourselves in it, we still stand within it. Now when the hindrance of sensible intuition is once removed, then we see ourselves in this spiritual community, and this is the other world; now these are not other things, but rather the same ones, but which we intuit differently. (28:289-9. Metaphysik L1 – my emphasis)¹²⁴

Here, roughly ten years after writing *Dreams* Kant makes it quite clear that he does not regard Swedenborg as having subjected immaterial substances to the conditions of sensibility.¹²⁵ Indeed, Kant actually credits Swedenborg himself with having postulated

¹²⁴ See also Kant's metaphysics lectures from 1792-3. Here Kant argues that "the concept of the spiritual life of the soul is wholly idea. It may be supposed; and if [after death] we pass over from the animal life into a purely spiritual life, then this is not to be sought in space. (Swedenborg assumed the ideal whole <totum> as real, invisible church.)" (*Metaphysik Dohna*, 28:690). Also suggesting that he did not regard Swedenborg as having offered an account of the spiritual life that is to be sought in space.

¹²⁵ Swedenborg himself argues that although angels are not "clothed with a material body" (*Heaven & Hell*, #77), they were once living human beings and are not "formless minds, nor ethereal gases, but people to a T" (*Heaven & Hell*, #75). In particular he will stress in *Heavenly Secrets* that spirits are not merely "abstract" Cartesian disembodied thinking subjects, but are essentially "organic" (*Heavenly Secrets*, p219).

the distinction between the intelligible and the sensible world.¹²⁶ Now, Laywine might argue that the Kant of the mid-1770s is misremembering the attitude of the Kant of the mid-1760s towards Swedenborg. However, even in *Dreams*, Kant makes it clear that he regards Swedenborg as having distinguished between spiritual ‘space’ and physical space. Spirits do have something analogous to positions, but these are not spatial positions. Thus Kant summarizes Swedenborg’s position in *Dreams*:

[T]he positions of the spirits, relative to each other, have nothing in common with the space of the corporeal world. Hence in what concerns their spirit-positions, the soul of someone in India may often be the closest neighbor of someone in Europe. (2:363)

Instead the relations and ‘distances’ between spirits are moral.

Their connections with each other are represented under the concomitant conditions of nearness, while their differences are represented as distances, just as the spirits themselves are not really extended, though they do present the appearance of human forms to each other. . . Everything depends on the relation of their inner state and on the connection which they have with each other, according to their agreement in the *true* and the *good*. (2:363)

This is an accurate characterization of Swedenborg’s position. For example, Swedenborg (1995) argues that, “there are no spaces in heaven except states that correspond to inner ones. . . Nearnesses are similarities, and distances dissimilarities. . . consequently, people who are in dissimilar [moral] states are far apart” (p. 192-3). This is why he believes that heavenly things “cannot be comprehended by a natural idea because there is space in that idea; for it is formed out of such things as are in the world; and in each and all things

¹²⁶ As we have seen in chapter one, Kant reaffirms this position in the corresponding sections of his metaphysics lectures throughout the 1780s. Although in these later lectures he does not mention Swedenborg by name the reference seems clear.

which strike the eye there is space".¹²⁷ Similar passages are extremely easy to find, and Kant obviously found them. It is difficult to understand how Laywine and Schönfeld could reach the conclusion that Swedenborg believed that immaterial things were subject to the spatio-temporal conditions of sensibility, or that that this is how Kant read him.

Swedenborg is insistent that angels neither exist in space nor experience heaven in spatio-temporal terms. He does believe that *immediately* after death existence often continues as it did on earth, and he tells a number of stories about dead spirits he met who did not realize that they were dead. Laywine takes his claim that some spirits after death do not recognize they are dead and experience the spirit world as if it is spatio-temporal as proof that he believes the world of spirits is spatio-temporal. She fails to recognize, however, that Swedenborg distinguishes between the life of the spirit immediately after death and the heavenly spirits in general and heaven as a particular community of spirits. Swedenborg makes this distinction because he believes that (some?) individuals need to make moral progress even after death, however he believes that at some point virtuous spirits will become angels and their inner eyes will be opened, and they will no longer experience the community they become part of as subject to the conditions of outer sense.¹²⁸

Although aspects of Kant's conception of the intelligible world can be traced back to Swedenborg, there are some significant differences. Most importantly, Kant objects to

¹²⁷ Quoted from Raine (1985), p.92.

¹²⁸ Swedenborg distinguishes between his visions of spirits and normal sensory vision. Sensory vision occurs by means of the sense organs, we experience things that strike our eyes. Swedenborg claims that his spiritual visions do not occur through his physical eyes, but through an inner eye.

Swedenborg's claim that objects in the phenomenal world can be symbols of the intelligible world of spirits. Thus he claims in his *Anthropology* that,

To claim that the actual phenomena of the world, which present themselves to the senses, are merely a symbol of an intelligible world hidden in the background (as Swedenborg does), is fanaticism. However, in the exhibition of concepts (called ideas) which belong to that morality which is the essence of all religion and which consequently come from pure reason, we must distinguish the outer shell, useful and necessary for a time, from the thing itself, the symbolic from the intellectual (public worship from religion) – this is enlightenment. If this is not done an ideal (of pure practical reason) would be replaced by an idol and the final purpose would be unsuccessful. (7:191-2)

Kant here objects to regarding the phenomenal world as a symbol of the spiritual world. What he is objecting to here is Swedenborg's doctrine of correspondences. For this would suggest that the phenomenal world was in need of (inspired) interpretation, which Kant finds morally problematic. Instead, Kant thinks that we have the pure idea of a spiritual world and we can, and should, *think* of the phenomenal world as a world of spirits (or autonomous individuals). This is not a matter of interpretation, but a matter of application (of an idea to an object of experience) and this application is a matter of choice and does not require any interpretation.

In addition to criticizing symbolic (spiritual) interpretations of the phenomenal world, in this passage Kant also comments upon the usefulness of symbolic representations of the intelligible. He suggests that such representations may be necessary for a time, but ultimately we must replace our symbolic representation of the intelligible world with an intellectual one. I suggest that here Kant is talking from personal experience, because, influenced by Swedenborg's writings, he first started to think of the intelligible world in

symbolic terms, however, as we shall see, he gradually came to intellectualize his conception of this world, thinking of it more and more as an idea of pure reason.

To Conclude: Kant realized that just because space is necessarily subject to the category of community, not every community is spatial. Indeed our pure idea of a community (the kingdom of ends, or an intelligible world) is the sort of thing that could never appear in space. To put it crudely, what Kant took from Swedenborg was the idea that relations didn't have to be spatial. In addition to believing that the spiritual community is not spatial, although it contains qualitative moral relations analogous to quantities spatial relations, Swedenborg conceives of the spirit world as governed by non-physical pneumatic laws. It is no coincidence, then, that at the time of reading Swedenborg Kant began to conceptualize the intelligible world as a community governed by non-physical laws.

I agree with Schönfeld (2001) that the pre-critical Kant believed that, "bodies and souls, or material and immaterial substances, are subject to the same laws" (p.244), and that this made it impossible for him to conceptualize a disembodied post mortem condition. However, upon reading Swedenborg he did not encounter a parody of his own earlier position, but an alternative to it, for Swedenborg clearly distinguishes between physical laws and spiritual (or what he calls pneumatic) laws. Kant clearly found the idea of a spiritual community governed by spiritual laws morally appealing, and the genesis of his moral ideal of a kingdom of ends can be traced back to this idea. However, by the 1780's Kant had come to see that in conceiving of a community of spirits it is not enough to

conceive of it as governed by pneumatic laws, but had come to see that these laws must be given by the members of the community itself. In other words, he reached the conclusion that we can only conceive of a community of spirits if we think of each individual spirit as autonomous. This is a notion that is not to be found in Swedenborg, and will be the theme of the following chapter.

Chapter Four

From Pneumatic Laws to Moral Laws:

Interaction and Autonomy

I have argued that the idea of a ‘good will’ should be understood as the idea of ‘a member of an intelligible world’, and that the idea of an intelligible world is the idea of a community of individuals in real interaction. In the previous chapter I argued that Kant first started to conceive of the intelligible world in these terms in the course of his engagement with Swedenborg. In this chapter I shall examine the development of this idea of a community in more detail. At the time of writing *Dreams* Kant conceived of such a community as a community of individuals governed by pneumatic (spiritual laws), and does not seem to have thought it to be important that these laws must be thought of as given by the members of the community. The mature Kant, however, conceives of such a community as governed by moral laws and believes that the only type of individuals that can really be members of a community are autonomous agents.

The reason for this is that an individual can only be a member of a unified (intelligible) world if the individual itself is the ground (or more precisely if it is the concurrent ground) of the intelligible relations between individuals that constitute the world, because for a world to be *a* world it must be unified and the principle (or source) of its unity must

be intrinsic to the world. And this is only possible if the individuals that constitute the world are the source of the unity of the world. Now, what unifies a world, and makes the world *a* world, are the relations (or laws) that hold between its individual members. So for a world really to be *a* world the individual members of the world must be the source of the laws that provide the world with its unity. Thus the idea of a member of *a* world is the idea of a being who “gives”, or legislates, the laws of the world. As a consequence, the only type of being that has the capacity to be (or become) a member of a ‘world’, strictly speaking, is a being that is aware of potential laws and chooses to actualize these laws, and this is, by definition, an autonomous agent.

In so far as it is possible for me to instantiate the moral ideal and become a member of an intelligible world, then, I would have to (a) be aware of the potential laws which could ‘glue’ such a world together, and (b) choose to actualize these potential laws. For only a being that has chosen to actualize these potential laws can be a member of an intelligible world. Thus, to choose to be a member of an intelligible world is to choose to be a legislator for such a world. Now, a being that has chosen to actualize law is, by definition, an autonomous agent. My consciousness of moral laws, then, is a consciousness of potential laws, which, if I choose to ‘give’ them, would give intelligible unity to the world. In other words, my consciousness of moral laws is a consciousness of the potential relations between individuals that could hold us together in one community. For me to become a member of the intelligible world is for me to choose to actualize the potential laws of this world, laws which if actualized would transform me from, at most, one of a multitude into a member of a community.

(4a) Pneumatic Laws and Forces in *Dreams*

In *Dreams* Kant muses on the possibility of ‘pneumatic laws’ (that is, laws governing spirits) through which I could be “connected” to spiritual beings “independently of the mediation of matter”. He writes that,

I am connected with beings of my own kind through the mediation of corporeal laws, but I can in no wise establish from what is given to me whether, in addition, I am not also connected, or could not ever be connected with such beings, in accordance with other laws, which I shall call pneumatic laws, and be so independently of the mediation of matter.
(2:370)

Here Kant expresses his skepticism as to any possible knowledge of the world of spirits. Experience can provide me with no knowledge of a world of spirits and of my possible relationship to such a world. Such skepticism, however, does not rule out the possibility of such a world of spirits connected by pneumatic laws, nor does it rule out the possibility that I am actually (or could possibly become) a member of such a world.¹²⁹

Kant makes the relationship between pneumatic laws, interaction and morality, clearer when he suggests that these putative ‘pneumatic laws’ between spirits can be thought of as analogous to the law of gravity. For just as the law of gravity is thought of as the basis for interaction between bits of matter, pneumatic laws can be thought of as the basis for interaction between spirits. In addition, just as the postulation of the law of gravity is the

¹²⁹ It is worth noting that even at this early stage in his development, Kant is not conceiving of the ‘spiritual world’ as a world that is ontologically distinct from the physical (or phenomenal) world. Kant does not suggest that the world of spirits is another world – but that ‘spiritual interaction’ (if possible) would merely be another way for me to be “connected with beings of my own Kind”.

basis of the postulation of gravitational forces, the postulation of pneumatic laws allows us to postulate pneumatic forces, and, having postulated such pneumatic forces we are free to regard our moral impulses as manifestations of these forces. Thus, Kant argues that Newton,

Did not hesitate to treat gravitation as a genuine effect produced by the universal activity of matter operating on itself; for this reason he also gave it the name 'attraction'. Are we, then, to suppose that it would not in the same way be possible to represent the phenomena of the moral impulses in thinking natures, who are reciprocally related to each other, as the effect of a genuinely active force, in virtue of which spirit-natures exercise an influence on each other? (2:335)

Here Kant argues that just as there was nothing illegitimate in Newton's introduction of the idea of attraction as a force to explain the law of gravity, there is nothing illegitimate in thinking of putative pneumatic laws as being grounded in 'pneumatic forces' and he suggests that this is a possible way for us to conceive of our own (actual) 'moral impulses'. In other words, Kant suggesting a way for us to think of our moral impulses as the basis of some sort of spiritual (or intelligible) interaction, presumably with other moral agents.

Kant has something like the following in mind: Our actual "moral impulses" are immediate objects of consciousness. We are immediately aware of certain of our 'inclinations'¹³⁰ as moral inclinations. We are not, however, immediately aware of such inclinations as forces. The idea of a pneumatic law, however, is not self-contradictory and so it is possible to think of myself as interacting with others "of my kind" according to

¹³⁰ And here I am not using 'inclination' in a technical sense. In his later moral philosophy 'inclination' will acquire a technical sense.

such laws. In addition, one can draw an analogy between the notion of the law of gravity and these putative pneumatic laws, and just as the existence of the law of gravity suggests the existence of gravitational forces (of attraction), the existence of pneumatic laws would seem to imply the existence of pneumatic forces. I have no cognition of these putative pneumatic laws. If, however, I am to think of myself as in real interaction with “other beings of my kind”, but “independently of the mediation of matter” (and I am able to think of myself in this way) I do so by thinking of myself as possessing some genuine active (pneumatic) forces. Thus, although I possess neither knowledge of any pneumatic laws nor direct awareness of any pneumatic forces there is nothing contradictory involved in thinking of my ‘moral impulses’ as phenomenal representations of such (albeit merely postulated) forces. Thus, although I do not *intuit* my moral inclinations as genuinely active forces (and so the awareness of these impulses can provide no theoretical evidence for the existence of pneumatic laws), I can *think* of these ‘moral impulses as the “effects of genuinely active forces”, insofar as I think of myself as a member of a world of spirits governed by pneumatic laws. In this way I can *think* of (although not *intuit*) myself as a center of intelligible forces and, as such, as the subject of intelligible interaction.

Expressed in the language of morality, we could say that it is the postulated existence of pneumatic laws that makes it possible for us to think of our wills as morally effective. For our moral impulses can only really be considered as forces (and hence as genuinely effective) if there were pneumatic laws. Thus Kant remarks on the following page that,

All the morality of actions, while never having its full effect in the corporeal life of man according to the order of nature, may well do so in the spirit-world according to pneumatic laws. (2:336)

Here, Kant suggest that if there were pneumatic laws, this would explain how the morality of actions could be effective. Combining this with the suggestions from the passages previously cited, we could say that it is through the postulation of pneumatic laws that we are able to think of our moral impulses as effective. To be effective is to have causal power. So Kant is suggesting that we can think of ourselves as spirits acting upon one another through the exercise of moral forces if we think of ourselves as members of a spiritual world governed by pneumatic laws.

Kant is, in effect, playing with the idea of a ‘Newtonian ethics’, with moral agents conceived of as centers of moral gravity. In *Dreams* he suggests that we can think of our moral *impulses* as representations of pneumatic forces. By the 1780s our consciousness of moral impulses will be replaced by our consciousness of moral imperatives (and laws) for the mature Kant will suggest that it is our consciousness of the categorical imperative which allows us to think of ourselves as centers of moral gravity and hence as capable of moral (intelligible) interaction. This is what he means when he argues in the *Critique of Practical Reason* that, “the moral law is the *ratio cognoscendi* of freedom [autonomy?]” (5:5).

Kant was taken by this analogy between the law of gravity and ‘pneumatic laws’. And this analogy is not merely an off the cuff remark made in the 1760s and then forgotten. Indeed, Kant returns to this analogy between the idea of Newton’s law of gravity and the

idea of the moral law at least twice almost thirty years later in the *Metaphysics of Morals*.

Thus, in Section 24 of the *Doctrine of the Elements of Ethics*, Kant writes,

In speaking of laws of duty (not laws of nature) and, among these, of laws for human beings' external relations with one another, we consider ourselves in a moral (intelligible) world where, by analogy with the physical world, attraction and repulsion bind together rational beings (on earth). The principle of mutual love admonishes them constantly to come closer to one another; that of the respect they owe one another, to keep themselves at a distance from one another; and should one of these great moral forces fail, "then nothingness (immorality), with gaping throat, would drink up the whole kingdom of (moral) beings like a drop of water" (if I may use Haller's words, but in a different reference). (6:449)¹³¹

Here Kant suggests that it is the 'laws of duty' which allow us to consider ourselves 'as in an intelligible world'. These laws can be thought of as 'binding together rational beings' in the same way that the laws of gravity binds together the physical world. Thus Kant clearly implies that it is this laws of duty which provide the intelligible world with its unity. Kant also returns to this analogy in the *Religion*. Comparing divine mysteries to the idea of gravity he writes that,

The cause of the universal gravity of all matter in the world is equally unknown to us, so much so that we can see that we shall never have cognition of it, since its very concept presupposes a first motive force unconditionally residing within it. Yet gravity is not a mystery; it can be made manifest to everyone, since its *law* is sufficiently cognized. When Newton represents it as if it were the divine presence in appearance (*omnipraesentia phaenomenon*), this is not an attempt to explain it (for the existence of God in space involves a contradiction) but a sublime analogy in which the mere union of corporeal beings into a cosmic whole is being visualized, in that an incorporeal cause is put beneath them – and so too would fare the attempt to comprehend the self-sufficient principle of the union of rational beings in the world into an ethical state, and to explain this union from that principle. We recognize only the duty that draws us to it. (6:138)

¹³¹ See also, 6:470

(4b) From Pneumatic Laws to Moral Laws

By the 1780s (that is, by the decade in which he wrote the first two *Critiques* and the *Groundwork*) Kant had come to see that the existence of (pneumatic) laws could only provide the intelligible world with an ideal unity. Only moral laws, that is laws given by the members of the world themselves could provide the world with real intrinsic unity. Because of this, Kant soon came to see that the existence of pneumatic laws would not be sufficient to explain real interaction. Indeed, after the 1760s Kant does not often refer to *pneumatic* laws. He does, however, frequently talk of *moral* laws and I suggest that in his later works the function of conceptualizing the possibility of spiritual interaction is taken over by concept of moral laws. The main conceptual distinction between the pneumatic laws Kant discusses in *Dreams* and the moral laws of his mature ethics has to do with their objectivity. Pneumatic laws, insofar as they are modeled on the law of gravity are objective in a strong sense. Moral laws, on the other hand, although they are objective in the sense of being *universal*, also have a subjective element in that they only exist as actual laws if the individuals subject to them choose to subject themselves to them. Although their possibility is objective, their actuality is subjective. A good way to think of this difference between these two types of laws is to appeal to the distinction between the ‘ground of actuality’ and the ‘ground of possibility’. As Kant explains this distinction in his lectures on metaphysics,

What contains the ground of actuality is called cause <*causa*> or principle of becoming <*principium fiendi*>; what contains the ground of possibility is called the principle of being <*principium essendi*>. (28:572)

A bit of matter attracted to another bit of matter is subject to the universal law of gravitation. The bit of matter, however, is neither the cause of the possibility nor of the actuality of the law of gravitation. In so far as Kant's putative 'pneumatic laws' of 1765 are thought of as analogous to the law of gravitation, the same could be said of them. This is not, however, how we think of the relationship between moral agents and moral laws. Kant believes that moral laws, like physical laws, are objective in the sense that the moral agent who is subject to such laws is not the ground of their possibility. We cannot, and do not, *create* the moral law *ex nihilo*. Rather we *recognize* it as a command and as a command that holds not just for us but for all rational beings. Thus Kant claims in the Religion that,

it is our universal human duty to *elevate* ourselves to [the] ideal of moral perfection, i.e. to the prototype of moral disposition in its entire purity, and for this very idea, which is presented to us by reason for emulation, can give us force. But precisely because we are not its authors but the idea has established itself in the human being without our comprehending how human nature could have even been receptive of it, it is better to say that that *prototype* has *come down* to us from heaven. (6:61)

Moral laws are objective in the sense of being not created by us and in their *universality*. Although moral laws are objective in the sense of being universal, they are subjective in the sense that they are only potential laws, and moral agents must be the ground of their actuality. In other words it is up to human beings to choose, through their actions, whether or not the moral law actually comes into existence as a law. Kant's mature theory of real interaction, then, doesn't merely require that individuals be subject to laws but that they must also be the 'ground of actuality' of the laws they are subject to. Given

the fact that, by definition, only autonomous beings are subject to laws they ‘give themselves’ in this sense, only autonomous beings can really be in interaction. Unlike the law of gravity or a putative ‘pneumatic’ law, then, a ‘moral’ law is merely a potential law (of nature), which it is up to the agent herself to actualize.

At the very least, it is clear that Kant slowly modified his account of the role law plays in interaction between the 1760s and 1780s. In the *Metaphysik Mrongovius* lectures of 1782-3 Kant could argue that,

Physical influence happens according to general laws, but the two systems of ideal connection [i.e., occasionalism and pre-established harmony] do not. (29:868)

Here Kant seems to suggest that the existence of general laws alone is sufficient to distinguish ‘real’ from ‘ideal’ interaction. By the early 1790s, however, Kant is arguing quite explicitly that being subject to general laws is not enough for there to be real interaction. Thus in the *Metaphysik Dohna* Lectures (1792-3) he argues that:

The way we represent substances in the phenomenal world, all dispute ceases, for space already brings them into interaction <*commercium*>. But if we think a world merely through the understanding, this is more difficult. The relation of many substances among one another according to general laws is called harmony, this is *without interaction* <*absque commercio*> if no concept of cause and effect between them takes place. In so far as they actually stand in real interaction, the system of substances can be called a system of physical influence. (28:666)

Here Kant claims, apparently contradicting the position he held at the time of the *Metaphysik Mrongovius*, that the existence of general laws, although it is necessary for real interaction is not sufficient. The reason for this is because by the 1780s Kant had

come to distinguish between an *ideal* and a *real* community and had reached the conclusion that real interaction is only possible between individuals if they constitute a real community. A real community is only possible if the individuals that constitute the community are the source of the unity of the community, and individuals can be the source of the unity of a community if they are the source of the laws that provide the community with its unity. In other words, a real community can only be a community of (law-giving) autonomous agents.

(4c) Moral Laws and Autonomy

In the *New Elucidation* (1755) Kant had clearly not adequately distinguished between the idea of a real and ideal whole and could argue that:

since the reciprocal relation between [substances] does not follow from the fact that God establishes simply their existence, unless the same schema of the divine understanding that gives them existence also sets up their relation insofar as it **represents** their existence as standing in reciprocity, **it is completely clear that the general connection of all things is due merely to this divine representation.** (1:413 – my emphasis)

Here Kant suggests that the connection between things can be explained in terms of a representation in the divine understanding. On this early account then the relations between individuals ultimately exist merely in the divine ‘representation’. The relations are contained in the ‘divine representation’, but it is not clear how or why this would mean that there really were relations between the substances themselves. Kant quickly came to see the inadequacy of this aspect of his earlier account and in his later mature terminology he would call such an account of interaction merely ‘ideal’ rather than of

‘real’ interaction. By 1782, Kant has clarified the distinction between ‘real’ and ‘ideal’ interaction and defines ‘ideal’ interaction in the following way:

The ideal connection <*nexus idealis*> is not connection in the things themselves, but rather merely in the idea of the observer who considers them. (*Metaphysic Mrongovius*, 29:866)

Here Kant clearly rejects his earlier position, for given this account of the distinction between ‘ideal’ and ‘real’ connection, the model of interaction offered in the *New Elucidation*, with the relations existing merely in the ‘representation’ of God, is merely ‘ideal’. By the early 1780s Kant, then, had come to realize the inadequacy of his earlier account. For real interaction requires more than the fact that a set of individuals are subject to laws given by God, for real interaction is only possible in a world of autonomous agents self governed by moral laws. The difference between ‘moral’ laws and other laws is that, by definition, individuals subject to moral laws choose to subject themselves to such laws. And an individual who chooses to subject himself to moral laws is an autonomous agent. Thus, real interaction requires the existence of individuals that have the *capacity* to be autonomous. Now, an individual that has the capacity to be autonomous must (a) be aware of the (potential) law and (b) be free to choose or will the law.

Kant believes that real interaction is only possible between autonomous agents (self) governed by the moral law because he is committed to the position that a world in which individuals ‘really’ interact would have to be a ‘real whole’. And he came to believe the a real whole could only exist if the members themselves were responsible for its unity. In

particular Kant was committed to the idea that in a real whole it must simultaneously be the case that (a) the whole be unified, with the ground for the unity of the whole being internal to the whole and (b) the individual members are ontologically prior to the whole. I will argue that Kant thought that we can only understand how such a real whole is possible by conceiving of its ‘members’ as autonomous agents.

Kant is explicit about this in the *Groundwork*. Here he defines an autonomous agent as a being that is “subject *only to laws given by himself but still universal*”. What he means by this is that an autonomous agent must be the ground of the actuality of the moral laws, but is not the ground of their possibility, for being universal the possibility of such laws must transcend the individual. He continues by explaining that the concept of autonomy leads to the idea of a kingdom of ends, which is depended upon it. In addition, he explains that,

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 *kingdom of ends*. // By a *kingdom* I understand a systematic union of various rational beings through common laws. (4:434)

I suggest that Kant here identifies the idea of a ‘kingdom of ends’ with the idea of an intelligible world, for, as we have seen the idea of an intelligible world is the idea of a systematic union of individuals.¹³² There are a number of places in which Kant seems to

¹³² For example, Kant argues in his metaphysic lectures that, “the aggregation of the substances in which there is no community still does not constitute a world. Reciprocal determination, the form of the world as a composite, rests upon the interaction. If we thought substances without real connection <*absque nexu reali*> and without interaction <*commercium*>, where every substance would have no community with one

identify a 'kingdom of ends' with an intelligible world. For example in *Groundwork II* itself he talks of "a world of rational beings (*mundus intelligibilis*) as a kingdom of ends" (4:438). And, in his lectures on ethics, given in 1785 around the time he was writing the *Groundwork*, Kant is able to argue that,

The autonomy of our will greatly elevates our worth. The **members of a kingdom of ends**, whose ruler is God, **are the true intellectual world**. Augustine and Leibniz called it the Kingdom of grace. In the realm of ends, God is supreme ruler; in the realm of nature, the ultimate cause. . . (*Ethik Mrongovius*, 29:629 – my emphasis)

If my identification the idea of a kingdom of ends with the idea of an intelligible world is correct, then what Kant is arguing in *Groundwork II* is that the idea of an intelligible world is somehow dependent upon the idea of an 'autonomous agent'. And I have argued that reason why the idea of an intelligible world is dependent upon the idea of an autonomous agent is because we can only conceive of individuals as interaction and as members of a whole which has intrinsic unity if we conceive of them as the source of the laws that provide to whole with its unity.

Kant provides further support for such an interpretation a few paragraphs later when he argues that,

For all rational beings stand under the *law* that each of them is to treat himself and all others *never merely as means* but always *at the same time as ends in themselves*. But from this there **arises a systematic union of rational beings through common objective laws**, that is, a kingdom, which can be called a kingdom of ends (admittedly only an ideal) because what these laws have as their purpose is just the relation of these beings to one another as ends and means. (4:434 - Kant's italics, my bolding)

another, then what would indeed be a multitude <*multitudo*>, but still not a world" (*Metaphysik L1*, 28:196).

Here Kant claims that moral laws are the source of the “systematic union” of rational beings. These laws are the source of the relations between individual members of the kingdom of ends, and as a result of this they are the source of the unity of such a kingdom. The idea of a systematic union of beings is the idea of a community, and here Kant makes it clear that the purpose of moral (“common objective”) laws is to provide the community of beings with its unity. These laws are what relate the members of the community to one another.

Kant is perhaps most explicit about the relationship between moral laws and community in his lectures. Here he claims that,

Morality, through which a system of all ends is possible, gives the rational creature a worth in and for itself by making it a member of this great realm of all ends. The **possibility of such a universal system of all ends is dependent solely on morality alone. For it is only insofar as all rational creatures act according to these eternal laws of reason that they can stand under a principle of community and together constitute a system of ends.** For example, if all human beings speak the truth, then among them a system of ends is possible; but if only one should lie, then his end is no longer in connection with the others. Hence the universal rule for judging the morality of an action is always this: If all human beings did this, could there still be a connection of ends. (28:1100 – my bolding, Kant's italics)

Here Kant is quite clear that he believes that only moral agents (autonomous agents) can “stand under a principle of community”.

(4d) Why Did Kant Draw the Conclusion that Real Interaction is Only Possible Between Autonomous Agents?

Although there is good textual evidence to support the claim that Kant believed that a real community must consist of autonomous agents, he is not explicit as to why he draws this conclusion. There are, however, two arguments which, I believe, might lie behind his position. Before entering into a discussion of the concept and idea of community, I will briefly examine these two arguments. The first argument is based upon assumptions about the concept of individuality, the second on assumptions about what is involved in the notion of a 'law'.

(1) The most plausible reconstruction of Kant's own motivation for drawing the conclusion that a real community must consist of autonomous agents is provided by thinking about what is involved in the idea of an individual, for Kant maintains that our idea of a real community is the idea of a composite individual, that is, an individual composed of individuals. This argument appeals to a **Complete Entity Principle of Individuation**. The principle was defended by Leibniz in his early *Disputatio Metaphysica de Principio Individui* of 1663, and I believe that Leibniz remained committed to such a principle throughout his career. Leibniz explains (and attempts to defend) this principle in the following terms:

That by means of which something is, by means of it that something is one in number. But any thing is by means of its entity. Therefore, [any thing is one in number by reason of its entity].¹³³

According to the complete entity principle of individuation, then, an individual must be the source of its own unity. This principle of individuation is to be distinguished from principles that place the source of an individual's unity to a part or aspect of the

¹³³ Leibniz (1996), *Disputatio*, #5.

individual, or to something external to the individual, for example, a theory that identifies a form or universal as the source of the unity of a material individual.

For Kant, then, the idea of a real community is the idea of a particular type of individual, namely a composite individual. To understand why Kant believed that such an individual must be composed of autonomous agents, it is instructive to examine Leibniz' failed attempts to conceptualize composition. As is well known, Leibniz believes that everything that exists is one. According to him, the only things that really exist are either themselves individuals or are composed of individuals. Leibniz, however, has a problem with explaining the possibility of composite individuals (or substances), for a composite substance must be composed of individuals and it must be an individual itself. For Kant a real community is composed of individuals and it is itself an individual, not necessarily in the sense of being an agent but at least in the sense of being unified.¹³⁴ I suggest that Leibniz must ultimately conclude that there can be no real composite individuals, for he is implicitly committed to the position that all composition is ideal.

The reason Leibniz felt compelled to deny the possibility of composite individuals was because he was committed to a particular principle of individuation, namely that "every

¹³⁴ Although, as we shall see when we look at Kant's account of property in the *Metaphysics of Morals*, Kant does seem to suggest that the type of community necessary to conceptualize the transfer of property is the idea of an agent. Here Kant is in a very Rousseauian mood and seems to be thinking of the moral community as a common will. In what follows, however, I am merely arguing that a real community must be an individual in the sense of being intrinsically unified and will not suggest that it must be thought of as an individual in the sense of being itself an agent.

individual is individuated by its complete entity”.¹³⁵ We can make sense of Kant’s position if we regard him as implicitly committed to the same (or at least a very similar) principle of individuation. Kant, however, believes that even with such a commitment one can make sense of the idea of a world or community of individuals. However, given a complete entity principle of individuation, such a community or world can only be a real community if the individuals which make up the community are themselves responsible for the unity of the community, and Kant thinks that this is possible if we think of the community as unified by laws and we think of the individual members of the community as the source of the laws (in the sense of being law-givers.) By assuming that Kant was committed to something like a complete entity principle of individuation, we can understand why he believed that such a community could only consist of autonomous agents. For “that by means of which a community is” are the individuals which make up the community, and if we assume, with Leibniz, that “that by means of which something is, by means of it that something is one in number”, then we must conclude that if a community really is unified then the only way this is possible is if the individuals which constitute the community are the source of its unity. Now, Kant believes that the form of a unified community are laws, and so if the individuals that make up a real community are to be thought of as the source of the unity of the community, they must be thought of as the source of its laws.

¹³⁵Leibniz (1996). This claim is from Leibniz’ *Disputatio Metaphysica de Principio Individui*. This is one of Leibniz’s earliest writings, dating back to 1663, when he was seventeen years old. I believe that his commitment to this principle of individuation remained solid throughout his philosophical career.

(2) Although I believe that an appeal to the Complete Entity Principle of Individuation provides the most plausible reconstruction of Kant's own motivation, I believe that a more plausible justification of Kant's position can be given, by thinking about what is involved in the idea of a law, and in particular, what is involved in the idea of being 'law governed'. This argument is influenced by H.L.A. Hart's analysis of law in *The Concept of Law*.

The basic point can be made in terms of the distinction between being 'subsumable under laws' and being 'law governed.' An observer of a multitude of substances could experience regularities in the relations between states of particular individuals and subsume these regularities under universal 'laws', but this does not in itself imply that there is any real relationship between the individuals themselves. The laws in such a case do in some sense 'unify' the individuals, but in such a case the unity is merely in the eyes of the observer. Real interaction involves not merely that there are observable regularities which can be **subsumed under laws**, but that the individuals themselves are **law-governed**. As H.L.A. Hart (1994) argues in *The Concept of Law*, laws should be understood as rules and rules (by definition) must have an 'internal aspect'. He imagines an 'objective' observer of society, who makes judgments about the existence of social laws merely on the basis of observable external behavior and argues that,

If the observer really keeps austerly to this extreme external point of view and does not give any account of the manner in which members of the group who accept the rules view their own regular behavior, his description of their life cannot be in terms of rules at all. . . . Instead it will be in terms of observable regularities of conduct, predictions, probabilities, and signs. (p.90)

To be justified in claiming that there is a law (which Hart argues is a type of rule) governing the behavior of a set of individuals, the actions, or behavior, of the individuals must not merely be **subsumable under** the ‘rule’, but **the rule itself must provide a reason for the behavior.**

(4e) Conclusion

In this chapter I have argued that Kant thinks that we can only think of a real community insofar as we think of its ‘parts’ or members as autonomous agents who logically proceed and are independent of the whole but are collectively responsible for the unity of the whole, in the sense that the moral laws, which are given by each member of the community, are the ground of the unity of the whole. If my interpretation is correct, the idea of an autonomous agent is not primarily an ethical concept, but is an idea of reason that plays an essential role in our conception of a community (or intelligible world) that is intrinsically unified. I have argued that seeds of this position can be traced back to *Dreams*. In this early work, however, although Kant argues that our idea of an intelligible community of spirits is the idea of a community unified by laws, he has not yet reached the position that such a community can only be united by laws that are given by the members of the community. This is a position he will only reach in the 1780s. In the following chapter I will examine Kant's account of the logical basis of our idea of autonomy and its relation to the table of judgments.

Part Two

METAPHYSICS

Chapter Five

The Critical Kant & Traditional Metaphysical

Enquiry

In the first part of the dissertation I was primarily concerned with questions that are clearly ethical. In the following four chapters I will concentrate on examining questions that may, on the face of it, appear to be entirely theoretical questions. In particular I will concentrate on examining Kant's theoretical account of the idea of community and the related notion of interaction. Before moving onto this, however, in this chapter I will make some general remarks about the role and status of metaphysical speculation in Kant's mature philosophy.

I have claimed that the ethical ideal of a good will or autonomous agent is to be identified with the metaphysical idea of a member of an intelligible world. As a result I believe that we can understand many of Kant's ethical commitments by examining what he believes is involved in the idea of a world, and in particular, with what is involved in the notion of 'membership' in such a world. Many of these 'metaphysical' commitments, which play an essential role in his ethical system, can be traced back to his pre-critical period. For, although there was a radical break in Kant's understanding of the *status* of metaphysical

claims at some point in the 1770s, there is much continuity in the *content* of his pre-critical and post-critical metaphysics.

These claims may sound implausible to those readers raised on the assumption that Kant's critical attitude towards traditional metaphysics is purely negative and destructive. Kant himself, however, is clear about the importance of traditional metaphysical enquiry in ethics. For example, in the *Metaphysics of Morals* he claims that,

no moral principle is based, as people sometimes suppose, on any feeling whatsoever. Any such principle is really an obscurely thought metaphysics that is inherent in every human being because of his rational predisposition, as a teacher will readily grant if he experiments in questioning his pupil socratically about the imperative of duty and its application to moral appraisal of his actions. – The way **the teacher presents this (his technique) should not always be metaphysical nor his terms scholastic, unless he wants to train his pupil as a philosopher. But his thought must go all the way back to the elements of metaphysics, without which no certitude or purity can be expected in the doctrine of virtue,** nor indeed any moving force. (6:376 – my emphasis)

Here Kant goes as far as to suggest that the moral *philosopher* must be taught metaphysics in scholastic terms. The remainder of this chapter will be devoted to examining the role of traditional metaphysics in Kant's critical philosophy. I will argue that although the critical Kant believes that metaphysical speculation cannot provide us with knowledge of the world, such speculation can provide us with a clearer 'image' of what we can become.

(5a) Introduction

I have claimed that we can come to understand something about Kant's ethics by examining his metaphysics, and in particular by examining what he has to say about the idea of an intelligible world. At first sight this may seem to be a strange claim to make about Kant, for he repeatedly claims that we can know (or 'cognize') nothing of the intelligible world. When Kant claims that we can know nothing of the intelligible world he is not claiming that we can say nothing interesting or useful about our *idea* of an intelligible world. Instead of banishing all metaphysical speculation as meaningless, he merely wishes to clarify the epistemic import of such speculation. Such speculation cannot provide us with any knowledge. The reason for this is that pure thought, governed by the law of non-contradiction, provides us with no criterion of real possibility. The fact that a concept does not contain a contradiction merely implies that the concept is thinkable. The fact that a concept is thinkable, however, does not imply that there actually is, or even could be, an object corresponding to the concept. Thinkability, then, is not an adequate criterion for real possibility. This commitment is the basis for Kant's radical break with the rationalist tradition. For if pure thought provides us with no criteria for real possibility we must posit some other, distinct, faculty (the faculty of intuition) that can explain how we are able to make such judgments.

The belief that thinkability is not an adequate criterion of real possibility is the basis for Kant's distinction between thinking and cognition. Kant makes this distinction in his preface to the second edition of the *Critique of Pure Reason*, explaining that,

To **cognize** an object, it is required that I be able to prove its possibility (whether by the testimony of experience from its actuality or *a priori* through reason.) But I can think whatever I like, as long as I do not

contradict myself, i.e., as long as my concept is a possible thought, even if I cannot give any assurance whether or not there is a corresponding object somewhere within the sum total of all possibilities. (Bxxvi)

Traditionally metaphysicians had thought that they were, in effect, providing a description of the way the world really is ‘in itself’. Kant, as everyone knows, is famous for rejecting this conception of the role of metaphysics and arguing that our cognition is restricted to objects of possible intuition, with ‘intuition’ being defined as that (cognitive) faculty through which objects are given. And because Kant is committed to the belief that intuitability is the only criterion we possess for real possibility, he concludes that we can have no cognition (or knowledge) of anything that is (in principle) incapable of appearing in space and time, for it is a condition of our form of intuition that the only things that we (can possibly) experience are things that are able to appear in space and time. Understanding the conceptual relationship between two pure concepts or ideas, then, cannot provide us with knowledge or cognition, if the objects of these ideas cannot be given in experience, for we have no criteria for judging whether the objects of such concepts are actually possible. Kant believes, for example, that it is impossible to think of (the idea of) a world of individuals without thinking of (the idea of) God as the ground of the world, but this does not mean that we ‘know’ or ‘cognize’ that God is the ground of the world, for we do not know there is, or even could be, anything corresponding to our idea of a world.

The idea of an intelligible world is the idea of a world that can be thought but not intuited (at least by beings such as us whose form of intuition is spatio-temporal). As a

consequence the idea of an intelligible world is not a possible object of experience, and so we have no criteria for judging whether such a world is really possible. There is nothing contradictory in the idea, but non-contradiction, Kant believes, is not an adequate criterion of what Kant calls *real* possibility. As a result of this Kant believes that the idea of an intelligible world can provide us with no knowledge, for it does not provide us with an object of cognition.

(5b) Kant's Critique of Traditional Metaphysics

Like Kant, rationalist metaphysicians distinguished, in principle, between the notion of an object of *experience* and an object of *knowledge/cognition* and believed that there were objects of knowledge that could not possibly be objects of experience. Thus, philosophers such as Descartes and Leibniz would agree with Kant that we have ideas of objects, such as God, that cannot possibly be objects of *experience*, but they did not assume that this meant that such ideas could not provide us with any knowledge. For these philosophers, God was not a possible object of experience, but he is an object of knowledge. Indeed in the fifth *Meditation* Descartes famously argued that we can know that God exists merely by examining our idea of God. Leibniz criticized Descartes' argument, pointing out that it rests upon the hidden assumption that God is possible. Leibniz, however, believes that we can prove that God is possible if we can show that the idea of God does not contain a contradiction. Kant rejects this assumption and insists that the non-contradictoriness of an idea is not an adequate criterion for the real (*de re*) possibility of the object of an idea.

Kant's radical reconceptualisation of modality

Kant believes, then, that we can think of things, *without contradiction*, which might actually be impossible. Thus, for example, in his lectures on metaphysics, Kant remarks that,

The concept of spirit has nothing contradictory in the representation, but whether it is possible that such an immaterial being can exist, this cannot be comprehended... [For] there can nevertheless be objects impossible in themselves that are assumed as possible because their concept experiences no contradiction. (29:962 – *Metaphysik Vigilantius*)

Here Kant makes it clear that he believes that non-contradictoriness is not an adequate criterion of *de re* possibility. For we may possess the concept of a spirit, and this concept may contain no contradictions, but this does not imply that spirits are really possible. Thus, although the ideas of immaterial spirits and of God contain no contradiction, this gives us no ground for assuming that immaterial spirits or God are really possible. For although there may be nothing contradictory in the idea of a being that is both omnipotent and omniscient, there may be a *de re* impossibility involved in the combination of omnipotence and omniscience. The analogy Kant wishes to draw here is something like the following: we can have the concept of a triangle, but we only know that triangles are possible because a triangle can be given in intuition, that is, because we can construct a triangle corresponding to our concept in (the pure intuition of) space. We do not know, however, whether there is any form of intuition in which an actual spirit corresponding to our idea could be given. As a consequence, any argument that takes as a premise the

claim that spirits, say, are really possible because the idea of a spirit contains no contradiction cannot be sound. This is why no reasoning from ideas can provide us with any *knowledge*, for any such reasoning must begin with the ungrounded assumption that the putative object of the idea is really possible. Working out the *conceptual relationships* between, say, our ideas of spirit, world and God, however, does not involve any appeal to the real possibility of the (putative) objects of these ideas, and so is a legitimate form of inquiry, so long as we do not make the further assumption that such investigations can provide us with knowledge.

Lying behind this position is a radical reconceptualisation of modal properties. For Kant, modal categories are not properties of objects, but instead concern the relationship between concepts and objects. When we claim that ‘unicorns are possible’ we are not asserting something about the nature of unicorns. Instead, we are making the claim that there could be an object corresponding to our concept ‘unicorn’. A more Kantian way of putting this would be to say that to claim that ‘unicorns are possible’ is to claim that an object could be *given* corresponding to the concept ‘unicorn’. Such an account of modal claims rests upon distinguishing between the faculty through which concepts are thought (the understanding) and the faculty through which objects are given, and Kant calls the faculty through which objects are *given* the faculty of ‘intuition’. Kant’s radical reconceptualisation of modality, then, lies behind the distinction he draws between understanding and intuition.

This distinction (between the understanding and intuition), then, was not introduced primarily to explain the particular nature of actual human experience, but is necessary for more abstract reasons. Specifically, Kant believes that the distinction between a faculty of thinking and a faculty through which objects are given is necessary to conceptualize adequately what is going on in modal judgments.¹³⁶ For if we reject the position that the non-contradictoriness of a concept is an adequate criterion of real possibility, we need to appeal to some other, non-logical criterion for real possibility. Non-contradictoriness is merely a criterion for whether a concept can be thought. Real possibility has to do with whether an object can be given that corresponds to the concept, and this is a question that cannot (in principle) be answered merely by examining the concept. Pure logic alone cannot, in principle, answer questions about real possibility, nor, as a consequence, questions about existence. Questions about real possibility (and hence about being) cannot be answered by a pure understanding alone. Any being that can make the distinction between actual and merely possible objects must possess both a faculty through which objects are given (a faculty of intuition) and a faculty through which objects are thought (the understanding).

The (logical) possibility of other forms of intuition

¹³⁶ Kant makes this clear in the *Critique of Judgment*. Here he argues that: “For if two entirely heterogeneous elements were not required for the exercise of these faculties, understanding for concepts and sensible intuition for objects corresponding to them, then there would be no such distinction (between the possible and the actual). . . all of our distinction between the merely possible and the actual rests on the fact that the former signifies only the position of the representation of a thing with respect to our concept and, in general, our faculty for thinking, while the latter signifies the positing of the thing in itself (apart from this concept)” (5:402).

To claim that x is really possible is to claim that there could be an intuition corresponding to the concept x . Thus, in the *Critique of Pure Reason*, Kant argues that,

[T]he possibility of a thing can never be proved merely through the non-contradictoriness of a concept of it, but only by the vouching of it with an intuition corresponding to this concept. (B307)

Not violating the law of contradiction is for Kant, then, a necessary but not sufficient condition for real possibility. Intuitability, then, is the only positive (theoretical) criterion for real possibility. Now, because *we* can only intuit things in space and time, the only positive criteria *we* can have for real possibility is whether something could possibly be experienced spatio-temporally. And, given the fact that everything experienced in time is experienced as non-simple and conditioned, this implies that we are only justified in asserting real possibility of non-simple and conditioned beings. Kant, however, does not draw the conclusion that just because the (putative) object of an idea is the sort of thing that could not be experienced by us the object must be impossible. Drawing such a conclusion would be just as dogmatic as assuming that simply because an idea does not contain a contradiction its putative object is possible. For all we know there might be beings (for example, angels or God) which experience objects in very different ways than we do and could experience objects that we are unable to experience.

Thus, Kant argues in the *Critique of Pure Reason* that,

[I]f. . . I leave out all intuition, then there still remains the form of thinking, i.e., the way of determining an object for the manifold of a possible intuition. Hence to this extent the categories extend further than sensible intuition, since they think objects in general without seeing to the particular manner (of sensibility) in which they might be given. But they do not thereby determine a greater sphere of objects, since **one cannot**

assume that such objects can be given without presupposing that another kind of intuition than the sensible kind is possible, which, however, we are by no means justified in doing. (A254/B310 – my emphasis)

The categories, then, ‘extend further than sensible intuition’ in the sense that there are ideas that we can think without contradiction which cannot, in principle, be given in our experience. Just because we are incapable of experiencing objects which correspond to such ideas, we cannot conclude that such objects do not admit of being experienced, for there might be some being with some other kind of intuition which could experience such objects. Indeed, it is possible that there could be a being that could experience the objects we experience as necessarily spatio-temporal (and hence as necessarily conditioned) in some other way. It would, however, be dogmatic to presuppose that such an alternative kind of intuition is possible, so we must remain agnostic.

Although we have no (theoretical) justification for assuming that another form of intuition is possible we have no theoretical justification for assuming that such a form of intuition is impossible, for there is nothing contradictory in the idea of a non-spatio-temporal form of intuition. The notion, then, of such a form of intuition is problematic. Kant believes, however, that we have important moral reasons for hoping that such a form of intuition is possible. Indeed, Kant himself often argues in his lectures that we must hope that after our (physical) deaths *our* way of experiencing objects (our ‘form of

intuition') will change and that *we* will be able to intuit the (putative) objects of our pure ideas.¹³⁷

To conclude: The only positive criterion for real possibility is whether something can be a possible object of intuition. On Kant's view, concepts provide us with objects of thought. An object of thought is 'really possible' if there could be an intuition corresponding to it. Now, our form of intuition is spatio-temporal. Therefore the only criterion *we* have for real possibility is whether something could be an object of spatio-temporal intuition. However, just because an object cannot be intuited by us, this does not imply that such objects are impossible, for, as we have seen, Kant thinks that there is nothing contradictory in the idea of some other form of intuition. We do not know that other forms of intuition (say divine intuition, or angelic intuition or perhaps even our post-mortem form of intuition) are possible, but neither do we know that other forms of intuition are impossible, and so, theoretically, we must remain agnostic on this issue. The idea of a 'spirit' (that is the idea of 'a member of an intelligible world') is thinkable as it is not self-contradictory. We do not *know*, however, whether an intelligible world is really possible and so we do not *know* whether spirits are 'really' possible, for we have no criteria for judging whether an object could be given (intuited) corresponding to our idea of a spirit. Spirits would be really possible if they were possible objects of intuition

¹³⁷ For example, in his metaphysics lectures of 1782-3, lectures he gave between the publication of the first and second edition of the *Critique of Pure Reason*, Kant argues that, "Now we find ourselves already in the intelligible world, and each human being can count himself as belonging, according to the constitution of his manner of thinking, either to the society of the blessed or of the damned. He is now only not conscious of it, and after death he will become conscious of this society . . . We are now already conscious through reason of finding ourselves in an intelligible realm; **after death we will intuit and cognize it** and then we are in an entirely different world that, however, is altered only in form, namely, where we cognize things as they are in themselves" (*Metaphysik Mrongovius*, 29:919-20 – my emphasis). See also: 28:296, 28:445 & 28:593. This doctrine was discussed in chapter three.

by some being (say 'God' or an 'angel') with a different form of intuition from our own. But we have no conception of what such a form of intuition would be like, or even whether such a form of intuition is possible. Kant concludes, then, that the ideas of pure reason do not contain any contradictions and are therefore thinkable, but we just do not know whether the putative objects of these ideas are even possible. Kant labels such concepts or ideas 'problematic'.

(5c) The Positive Role of Traditional Metaphysical Speculation

To understand Kant's attitude towards metaphysics, and the positive role he believes it plays in ethical thinking, it is instructive to compare his attitude to that of Leibniz. Leibniz, in writing his *Monadology*, was attempting to provide a description of the way the world is in itself. Kant, in contrast, believes that we can have no knowledge of the 'way the world really is', no knowledge of 'things-in-themselves'. Although such metaphysical speculation can provide us with no knowledge of the way the world is, it can, however, provide us with an 'image' of the way the world could and should be. Central to Kant's critical philosophy, then, is the claim that we can have no knowledge of things as they are in themselves, and insofar as traditional metaphysics is understood as an attempt to give us such knowledge, Kant rejects it. This, however, does not mean he must reject all traditional metaphysical speculation, for such speculation may play an important role in moral reflection, in so far as it can provide us with a clearer understanding of an idea which has the power to determine our faculty of desire.

The intellect, the faculty of desire and the faculty of intuition

The Leibnizian rationalists, such as Wolff, argued that there is only one basic human faculty, the faculty of representation. Kant, following Crusius, insists that the faculty of representation (or the theoretical faculty) is distinct from the will, or what he calls the faculty of desire.¹³⁸ Understanding the distinction between the faculty of intuition and the faculty of desire, and the difference in the relationship of the intellect to these two faculties, is essential for understanding Kant's project, and in particular for understanding the role of traditional metaphysical questions in Kant's critical project.

Kant wrote three *Critiques* and each is concerned with the relationship between the intellect (which he often calls 'reason' or the 'understanding') and a particular 'faculty'. The *Critique of Pure Reason* deals with experience, which Kant believes involves the relationship of the intellect to the faculty of intuition, the *Critique of Practical Reason* is concerned with the relationship between the intellect and the will, or what Kant calls the faculty of desire, and the *Critique of Judgment* (or at least the first part, the *Critique of Aesthetic Judgment*) deals with the relationship between the intellect and our faculty of feeling (that is, the faculty of feeling pleasure and displeasure). Kant's account of the faculty of feeling is beyond the scope of my dissertation.

¹³⁸ And, as we have already seen, Kant also disagrees with Leibniz and Wolff in that he posits two distinct types of representation: intuitions and concepts, and corresponding to these two distinct faculties of representation: the faculty of intuition and the intellect.

As Kant explains this tri-partite division of philosophy in his lectures on metaphysics from the mid-1780s:

The faculty of knowledge, the faculty of feeling and the faculty of desire, are the three powers of the human soul. In all three, understanding and sense can come into play. If understanding is present, then the following sciences are possible: (1) Logic, in regard to the understanding; (2) aesthetic, the feeling of pleasure and displeasure in the understanding, which is taste; (3) practical philosophy, the faculty of desire in relation to the understanding. (*Metaphysik Mrongovius*, 29:597)

There are some problems with Kant's terminology, for he is not particularly consistent. He often uses 'reason', 'understanding' and 'intellect' interchangeably to refer to the intellectual faculty as a whole. However, at other times he uses 'reason' and 'understanding' in contrast to one another, using them to refer to particular aspects of the intellect. Thus, in the *Critique of Pure Reason* in particular, he often distinguishes between reason as the faculty of ideas and understanding as the faculty of concepts. Taken together they constitute the intellect. Wherever possible I will use the term 'intellect' to refer to the faculty as a whole, as it is the least ambiguous of Kant's terms. I am primarily interested in the relationship between the intellect and the faculty of desire.

As we have seen, for Kant, the faculty of intuition is the faculty through which objects are given to us. The faculty of desire, in contrast, is the faculty of choice. As Kant defines it, **“the faculty of desire is the faculty to be, by means of one's representations, the cause of the objects of these representations”** (*Metaphysics of Morals*, 6:211 – my emphasis).

Although Kant is committed to the position that the intellect (or pure reason) cannot provide us with **cognition**, he believes that it provides us with ideas. Amongst these ideas is the idea of a member of an intelligible world, and although this idea cannot be **an object of our faculty of intuition** it can be **an object of our faculty of desire (or the will)**. Thus, although Kant believes that the ‘intelligible world’ is not a possible object of intuition (and as a result not a possible object of cognition), it *is* a possible object of choice, for we can choose to be a member of such a world. To make the idea of an intelligible world the object of our faculty of desire is to make the choice that such a world actually exists, for to claim that it is possible to make this idea an object of our faculty of desire is, by definition, to claim that we can be the cause of the object of the idea.¹³⁹ As Sullivan (1989) explains, Kant believes that “moral reasoning has the unique power to bring into existence its own proper objects” (p.104).

The notion that through our choice we choose that the objects of certain pure ideas actually exist lies behind Kant’s claim in the *Critique of Practical Reason* that,

the upright man may well say: “I will that there be a God, that my existence in this world be also an existence in a pure world of the understanding, beyond natural connections, and finally that my duration be endless.” (5:143)

The upright man, Kant believes, chooses that his pure ideas have existence. Or, to say this in other words, the upright man chooses to give these ideas what Kant calls ‘objective reality’. Thus, although it is impossible for us ever to experience anything as

¹³⁹ For, “the faculty of desire is the faculty to be, by means of one’s representations, the cause of the objects of these representations”. If it is possible for us to make the idea of being a member of an intelligible world the object of our faculty of desire, then, by definition, by means of this representation (idea) of a member of an intelligible world, we can be the cause of the object of this idea.

an instantiation of the moral ideal, for an idea of pure reason is not a possible object of experience, it is possible for us to choose to be such an individual and in so doing to choose that an intelligible world actually exist. This is a consequence of Kant's definition of the faculty of desire as "the faculty to be, by means of one's representations, the cause of the objects of these representations." If we actually possess such a faculty and possess the idea of an intelligible world, to make this idea (or representation) the object of our faculty of desire would be, by definition, to be the cause of the object of this idea.

Now, just as our faculty of intuition has a certain subjective, human form, so does our faculty of desire. The subjective form of our faculty of intuition is space/time. This form is 'subjective' in that there is nothing contradictory in the thought of a being that intuits objects in some other, non-spatio-temporal, way. So we cannot say that it is a feature of a faculty of intuition, *as such*, that its form is to be able only to intuit objects in space and time, for we can still make some sense of the notion of a faculty of intuition even if we abstract from the spatio-temporal nature of our own intuition. The subjective form of our faculty of desire is to act on maxims.¹⁴⁰ Ideas of pure reason, then, cannot be objects of our faculty of intuition because our subjective form of intuition is spatio-temporal and objects corresponding to our ideas cannot be given in space-time. They can, however, be objects of our faculty of desire because the subjective form of our faculty of desire is to

¹⁴⁰ Kant defines it in the *Groundwork*, "[a] *maxim* is the subjective principle of volition" (4:401 – see also: 4:421), and I am suggesting that a maxim is a subjective principle of volition (i.e. a subjective principle of the faculty of desire) in a manner analogous to the way space and time are subjective forms of intuition.

act upon maxims, and an idea can be an object of a maxim, for I can make it my fundamental maxim to be a member of an intelligible world.

Metaphysical enquiry as conceptual clarification

Instead of providing us with knowledge, metaphysical speculation is ‘merely’ a process of concept-clarification. By asking questions about the nature of an intelligible world we are not discovering anything about being itself, but are merely drawing out and clarifying what is involved in our pure idea of a world. Now, if we believed that philosophy is solely concerned with providing us with knowledge, such conceptual clarification would seem to be a pretty pointless activity, at best a harmless conceptual game like building castles in the air. This is not, however, what Kant believes. For we are not merely spectators who take a purely theoretical interest in the world, but also actors who must make decisions about what to do and what type of person we want to be. And while our pure ideas cannot provide us with any theoretical knowledge, they can serve as ideals which can guide our actions and provide us with an ‘image’ of what we can choose to be. Kant is an ethical idealist and, as such, he believes that the choice to be moral is the choice to take a particular pure idea, namely that of ‘a member of the intelligible world’, as our practical ideal. In other words, Kant believes that we can make a pure idea the object our faculty of desire by making it the object of our fundamental practical maxim. He believes, then, that (pure theoretical) reason provides us with the idea of a particular kind of individual, namely the idea of ‘a member of an intelligible world’, or ‘spirit’. To be moral (a good will) is to instantiate this ideal. In Kant’s terminology, ideas of pure

reason are not given *objective reality* by the faculty of intuition; objects corresponding to ideas of pure reason can never be given in our experience. Kant believes, however, that they can be provided with objective reality by the will or faculty of desire. Thus he argues in the *Critique of Practical Reason* that the moral law,

must therefore be the idea of a nature not given empirically and yet possible through freedom, hence **a supersensible nature to which we give objective reality** at least in a practical respect, since we regard it as an object of our will as pure rational beings. (5:44 – my emphasis)

The “idea of a nature not given empirically”, supersensible nature, is our pure idea of a world. And in this passage Kant makes it clear that we can give this idea objective reality by making it the object of our will (or what he elsewhere calls our faculty of desire).

For Kant, then, to be moral is to take our pure idea of a world as the object of our faculty of desire. In so doing we can hope to give objective reality to the idea, and can hope to become a member of such a world. This is the point Kant is attempting to make in the *Critique of Practical Reason* when he argues that the moral law,

is to furnish the sensible world, as a sensible nature. . . , with the form of a world of the understanding, that is of a *supersensible* nature. . . . [S]upersensible nature, so far as we can make ourselves a concept of it, is nothing other than *a nature under the autonomy of pure practical reason*. The law of this autonomy, however, is the moral law, which is therefore the fundamental law of a supersensible nature and of a pure world of the understanding, the counterpart of which is to exist in the sensible world but without infringing upon its laws. The former could be called the *archetypal world (natura archetypa)* which we cognize only in reason, whereas the latter could be called the *ectypal world (natura ectypa)* because it contains the possible effect of the idea of the former as the determining ground of the will. For the moral law in fact transfers us, in idea, into a nature in which pure reason, if it were accompanied with suitable physical power, would produce the highest good, and it

determines our will to confer on the sensible world the form of a whole of rational beings.” (5:43)

Platonic idealism vs. ethical idealism

Having examined the distinction between the faculty of intuition and the faculty of desire, I am now in a position to explain the distinction Kant himself draws in the *Critique of Pure Reason* between Plato’s theoretical idealism and his own practical/moral idealism. According to Kant the platonic idealist believes that our ideas are archetypes of which the objects we experience are copies. In Kantian terms, a platonic idealist believes that ideas can be objects of (or determine) our faculty of intuition, whereas a moral idealist, such as Kant, believes that such ideas can only be objects for (or determine) our faculty of desire. Another way of putting this would be to claim that the platonic idealist believes that ideas can be instantiated in experience, whereas the moral idealist believes that ideas can only be instantiated in practice, that is, that we can practically instantiate these ideas. Understanding this distinction helps us understand Kant’s distinction between the theoretical and the practical, and in particular Kant’s belief that practical reason (or more precisely: reason in its practical employment) is creative in a way theoretical reason is not. This ‘creativity’ of reason in its practical use is central to Kant’s ethical idealism. Reason is creative practically in the sense that through our faculty of desire an idea, that is the idea of a member of an intelligible world, or (what amounts to the same thing) the idea of an autonomous agent, can be made real. The idea of an autonomous agent, then, is an idea of pure reason. As such nothing in experience can ever be adequate to it. That

is, we are unable to experience any object as an autonomous agent. We are, however, capable of choosing to be an autonomous agent and in making such a choice we can *hope* that we actually succeed in instantiating the idea. We have, Kant believes, the capacity to create a world, in the sense that the intelligible world can only exist if we (as its potential members) choose to create (or realize) it. In practice the choice we have is the choice to realize the potential laws that provide the world with its intelligible unity and have the capacity to make the world a world.

Kant makes the distinction between his type of idealism and Plato's in the section on the *Ideal of Pure Reason* in the *Critique of Pure Reason*, where he explains that,

[W]e have to admit that human reason contains not only ideas but also ideals, which do not, to be sure, have creative power like the *platonic* idea, but still have *practical* power (as regulative principles) grounding the possibility of the perfection of certain *actions*. (A569/B597)

Here Kant distinguishes between ideas and ideals and argues that whereas the platonic idealist thought that ideas have 'creative power' he believes that they only have 'practical power'. What Kant means here is that whereas the platonic idealist believes that ideas can be creative theoretically, he believes that they can only be creative (or, I would suggest 'effective') practically. The point Kant is trying to make here in distinguishing between platonic idealism and what I have called his 'ethical idealism' is the following. For the platonic idealist, ideas are creative in the sense that objects in the world (of experience) are (imperfect) copies or manifestations of ideas. According to platonic idealism ideas are understood as the 'ground' of the objects in the world of sense. For the platonic idealist, then, the objects we experience are copies (albeit imperfect copies) of ideas. Kant clearly

rejects this position. For Kant, ideas are (a) not the sort of things that could ever be experienced. And as a result of this fact, our ideas (b) cannot provide us with any theoretical knowledge. Although our pure *a priori* ideas cannot produce or create objects that can be experienced, these ideas can produce actions. And, more than this, our ideal of a good will can be practically effective in that we can choose to be a good will. The choice to be a good will would be an example of an ideal possessing creative practical power in a strong sense. For the existence of the idea (as an object of thought) is logically prior to (and a necessary condition for the existence of) the reality. Kant explains this practical ‘power’ or efficaciousness’ of ideals. He writes,

Thus just as the idea gives the rule, so the ideal in such a case [he has been talking about the Stoic ideal of the sage] serves as the original image for the thoroughgoing determination of the copy; and we have in us no other standard for our action than the conduct of this divine human being, with which we can compare ourselves, judging ourselves and thereby improving ourselves, even though we can never reach the standard. (A569/B597)

This is not a chance remark, and the claim that the only standard of our actions is provided by an ideal is central to Kant’s ethics. Here Kant suggests that the moral ideal should be understood as the ‘original image’ which can ‘thoroughly determine’ the copy. The platonic idealist maintains that ideas can determine our faculty of intuition. Kant rejects ‘platonic idealism’, believing that ideas cannot determine our faculty of intuition since our intuition is spatio-temporal and ideas cannot be instantiated in space/time. A particular pure ideal, the ideal of an autonomous agent, can, however, determine our faculty of desire, which, if we truly choose to be autonomous, can become a copy of this

original image.¹⁴¹ That is, Kant believes that it is possible for us to choose to be autonomous. In choosing to be autonomous we take the pure idea of an autonomous agent as the object of our faculty of desire. Of course, from the theoretical perspective we have no way of knowing whether we are (or have succeeded in being) autonomous. We cannot *know* whether we have actually chosen to be autonomous. When I look at myself I experience myself as existing in time, making a series of particular choices. It is possible, however, to think of a single and simple choice laying behind this series of particular choices. Kant calls this idea of a simple choice laying behind our series of particular phenomenal choices our intelligible character or disposition.¹⁴² At any particular time the most we can do is to ensure that our particular choices are compatible with the choices an autonomous agent would make. There is nothing contradictory, however, in thinking of a particular simple choice of a ‘fundamental project’ as the ground or source of all of our particular choices and actions, and hoping that this choice is the choice to be an autonomous agent.

¹⁴¹ In fact Kant’s account is more subtle than this. He believes that a will (faculty of desire) that was full determined by such an idea would be a *holy* will. He believes that we should strive for holiness but the most we can hope to achieve is to become virtuous. In his later work Kant was pre-occupied with the relationship between virtue and holiness. This topic is beyond the scope of my dissertation, but it is an important one, for insofar as reason can be practical we must, in some sense, hope that it is possible for us to become *holy*. I suggest that Kant’s considered position was that a virtuous will is a will that is *converging* on holiness. In the sense that the series $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16} \dots$ converges on one. We cannot complete the successive addition but we can ‘see’ that this series converges on one. Similarly, I believe, Kant believes that we can hope that there exists a divine intellect, which can see our virtuous striving for holiness as actually converging on holiness. Such an intellect, Kant hopes, would be able to take in the series of our actions in a ‘single glance’ and would see such a virtuous will as a holy one. We can, then, in this sense hope to instantiate the idea of a holy will.

¹⁴² In Sartre’s terminology, one could call this ‘disposition’, or ‘intelligible character’, one’s basic or original project.

The moral ideal as an object of rational cosmology

I have claimed that Kant is a moral idealist and have argued that the moral ideal is not the product of some distinct faculty of practical reason, but is an idea of theoretical reason. The question now is: upon which theoretical idea is our moral ideal based? In different contexts Kant uses different expressions to refer to this ideal. He talks, for example, of the idea of “a good will”, “an autonomous being”, “a member of the kingdom of ends” and “a divine human being”. All of these expressions refer to the same ideal, namely the idea of a member of the intelligible world. As such, the moral ideal must be understood as an idea of rational cosmology.

To understand this claim it is helpful to understand something about how metaphysics as a discipline was structured in 18th century Germany. Eighteenth century German metaphysics text-books divided metaphysics into general metaphysics (ontology) and special metaphysics. Special metaphysics was divided into three special sciences corresponding to the three objects of rational cognition, namely: rational psychology, rational cosmology and rational theology. Rational psychology was concerned with rational cognition of the **soul**; rational cosmology dealt with rational cognition of the **world**; and rational theology dealt with rational cognition of **God**. Although Kant rejected the possibility of rational cognition – that is cognition of objects through pure reason – the structure of the *Critique of Pure Reason* follows this traditional plan. Thus, although Kant rejects the possibility of ontology in the traditional sense, the first half of the *Critique of Pure Reason* can be understood as corresponding to the traditional role of

general metaphysics, although here, ontology, in the strict sense of the ‘science of being’, has been replaced by a ‘doctrine of elements’. For Kant believes that a science of being is not possible, for the intellect can give us no access to things in themselves.¹⁴³ As Kant explains in his lectures on metaphysics,

Ontology is a pure doctrine of all our *a priori* cognitions; or it contains the summation of all our pure concepts that *we* can have *a priori* of things.
(*Metaphysik L2*, 28:541)

According to the critical Kant, then, ontology cannot tell us anything about being or about things in themselves. Instead, it only provides us with information about our own cognitive capacities and faculties, and the contents of these faculties. In addition, although reason in its pure use cannot provide us with knowledge of objects, it can tell us something about its own limits. The *Transcendental Dialectic*, on the other hand, corresponds to the traditional disciplines of special metaphysics and is structured according to the traditional division of special metaphysics into three special sciences. Whereas traditional German metaphysicians understood these three special sciences as being concerned with three distinct types of object, which could be cognized by the human intellect, in Kant’s *Dialectic* these three ‘objects’ are merely objects of thought and can be examined purely as ideas. These ideas are possible as objects of our thought, but we have no way of knowing if there are, or even if there possibly could be, ‘real’ objects corresponding to them.¹⁴⁴

¹⁴³ For, having rejected the rationalist claim that non-contradiction is a sufficient criterion for real possibility, Kant believes that pure theoretical speculation can tell us nothing positive about the nature of being (or beings).

¹⁴⁴ As Kant explains in his lectures on metaphysics: “All intuition is sensible, and to cognise something without it would be, were it possible, supersensible cognition, which would also have to rest on supersensible intuition; all supersensible concepts [and here, I suggest, Kant is being imprecise for he is clearly talking about what he normally refers to as ‘**ideas** of pure reason’, and not to a priori **concepts**, such

Thus, in the *Critique of Pure Reason* Kant identifies three ideas of pure reason: the thinking subject, which is the object of **psychology**, the world, which is the object of **cosmology** and God, which is the object of **theology**.¹⁴⁵ The idea of a good will is an idea of pure reason and, as such, it must be identified with or derived from one of the three ideas of pure reason: the soul, the world or God.¹⁴⁶ And I argue that the idea of a good will (or autonomous agent) is an idea of rational cosmology, for it is the idea of a member of a world.

Evidence for this interpretation is provided by Kant's discussion of the ideas of practical reason in the *Dialectic of the Critique of Practical Reason*. Here Kant himself explicitly identifies the idea of autonomy (referring to it as 'the idea of freedom in the positive sense') as an object of rational cosmology. In sections VI and VII of the *Critique of Practical Reason*, Kant introduces three ideas 'of practical reason', corresponding to the

as the categories] are of the kind that even if otherwise one may well have grounds for thinking them, no intuitions can be given for them; thus they still lack an objective reality that would be connected with the concept, e.g., the concept of spirit" (*Metaphysik Vigilantius*, 29:967).

¹⁴⁵ And, as we have seen, these ideas are 'pure' in the sense that they have no sensible content. It is not merely that they have no empirical content (which would make them merely *a priori*) but they are ideas of pure reason, and as such, by definition, have no content provided by the faculty of intuition. As Kant explains, "A *pure* concept is one that is not abstracted from experience but arises rather from the understanding even as to content." (9:92) In addition, these ideas of pure reason, although they are possible objects of thought, are not possible objects of experience (or intuition) and as a consequence they are not possible objects of cognition.

¹⁴⁶ Ultimately I think that all three ideas play a role in our practical deliberation, for ultimately the three ideas involve each other. In fact one could say that there is ultimately only one (but complex) idea of pure reason: namely, the idea of a **world** (community) of **individuals** created by **God**. To be moral is to choose to be a member or part of such a community. We cannot, for example, think of the idea of a world of agents without thinking of God as the cause of such a world, for our idea of a community of moral agents is, Kant believes, the idea of a community created by God. This relationship between the idea of God and the idea of a world is purely conceptual and analytic. The choice to be moral is the choice to be 'a member of a (non spatio-temporal) world of agents created by God'. This is why Kant claims that the "upright man" wills that God exists. For if, upon analysis, we work out that our idea of a world can only be the idea of a world created by God, then to choose that there be a world is to choose that there be God. A defense of this claim, however, is beyond the scope of this dissertation.

three ideas of reason introduced in the Dialectic of the first *Critique*. These ideas are the ideas of ‘immortality’, ‘freedom’ and ‘God’. And, in this passage, when Kant talks of ‘freedom’ what he really means is the idea of an autonomous agent. In section VII Kant is attempting to explain what is necessary “in order to extend pure cognition *practically*” and what he means is that he is attempting to explain what must be the case if pure reason is to have a practical application. He argues that the practical extension of pure cognition,

is not possible without presupposing three theoretical concepts (for which, because they are only pure rational concepts, no corresponding intuition can be found and consequently, by the theoretical path, no objective reality): namely, freedom, immortality and God. (5:134)

The first interesting point about this passage is that Kant calls these three ideas of practical reason theoretical concepts, identifying them with the three ideas of pure reason identified in the *Critique of Pure Reason*. They correspond to the three ideas introduced in the *Critique of Pure Reason* in the following terms: The practical idea of immortality is a object of rational psychology, the practical idea of freedom/autonomy is an object of rational cosmology and the practical idea of God is, obviously, an object of rational theology. A few pages later Kant makes this identification explicit. Here Kant is, once again, discussing these three practical ideas, but now he labels them differently, for here he identifies them as “these ideas of God, of an intelligible world (the kingdom of God), and of immortality” (5:137). The idea of ‘freedom’ seems to have been replaced by the idea of an ‘intelligible world’. Kant seems to be using the idea of ‘freedom’ and the idea of ‘an intelligible world’ interchangeably. I suggest that the reason for this is because he believes that the idea of a free (or autonomous) agent *is* the idea of a member of an

intelligible world. For the idea of the intelligible world can only be the idea of a world of autonomous agents.

Indeed, Kant himself in this section defines “freedom considered positively” as “the causality of a being insofar as it belongs to the intelligible world” (5:132). We should take Kant literally here. The idea of freedom in the positive sense, that is, the idea of autonomy, *is* the idea of a being insofar as it belongs to the intelligible world. To be free in the positive sense **is** to be autonomous, **is** to be a member of an intelligible world. Kant makes it clear, then, that the idea of freedom or autonomy should be understood as an object of rational **cosmology**.

At first sight this may seem a very strange claim to make, for one would think that the idea of a ‘spirit’ or of an autonomous *agent* would surely be an object of psychology rather than cosmology, corresponding to the concept of the soul rather than somehow being derived from the idea of a world. It is clear, however, that the idea of an ‘autonomous agent’ or ‘good will’ is essentially the idea of a **member** of an intelligible **world**, and as a consequence an adequate understanding of what Kant means by a ‘good will’ requires an understanding of how Kant conceives of the idea of a ‘World’. Kant’s most detailed accounts of what is involved in the idea of a ‘world’ are to be found in his lectures on metaphysics, and in his pre-critical writings. Although the critical Kant radically changed his position on the status of the idea of the intelligible world, there is a continuity (although clearly also a development) between his pre-critical and post-critical understanding of the idea of the nature (or content) of the idea of the intelligible world.

For this reason I believe that paying attention to what Kant has to say about the nature of the intelligible world in his pre-critical writings can help us understand his critical position. A careful reading of these texts, then, is needed to fully understand Kant's ethical writings.

(5d) Kant's Account of the Phenomenal World and the Conditioned Nature of All Objects of Experience

The idea of an intelligible world and the idea of a member of an intelligible world, then, are ideas of pure reason. The (putative) objects of our ideas of pure reason are not possible objects of experience because they are not the sorts of things that we could ever come across in our experience. This is not a contingent, empirical claim but is based upon the difference in the intrinsic nature of a possible object of experience and an idea of pure reason, for ideas of pure reason are, by definition, objects of the pure intellect and they are unconditioned. Anything that is a possible object of experience, however, is conditioned. In this final section of the chapter I will briefly examine Kant's account of the nature of our phenomenal experience, explaining why he maintains that all objects of experience are conditioned. This fact about objects of experience is the reason why nothing can be given in experience corresponding to our (unconditioned) pure ideas.

The reason why all objects of experience are necessarily experienced as conditioned is because we can only experience objects as given in space and time, and an essential feature of the very structure of both space and time is that anything experienced in space and/or time is experienced *as* conditioned. Although Kant himself does not explicitly

present his argument in such general terms, I think the following type of argument lies behind his thought: The experience of the ‘presence’ of objects is a fundamental feature of temporal experience. To experience something in time, then, is to experience it as present, as existing now, in the present time. But a condition of experiencing something as present is the implicit awareness that there is a past, that is not experienced. This implicit awareness of a past that is not and cannot be experienced is a condition for experiencing anything present *as* present. In other words, the experience that something *is* involves an awareness (or what Kant calls a ‘representation’) that there *was* something prior to what is now being experienced. Everything that is experienced as in time, then, is experienced as having a past. Thus, everything experienced in time is experienced as conditioned, for it implies or points to the existence of something existing in some prior time that is not-present, and not experienced, but that is its condition for being (experienced) in time.

A similar argument can be made with regard to space. As Kant explains in his lectures on metaphysics, “No space, no time can be thought without at the same time thinking of a much larger space or time. . .” (*Metaphysik Vigilantius*, 29:977).¹⁴⁷ When we experience something as occupying a particular space we experience the space as bounded, for we are always implicitly aware that what is given in our visual field is only a part of what there is. This is the sense in which the experience of an object as occupying a particular space is conditioned by the thought of the ‘larger space’ to which the particular space belongs. This ‘larger space’ is not immediately experienced but the thought of some

¹⁴⁷ He makes a very similar point in the *Transcendental Aesthetic*.

'larger space' is a necessary condition for the experience of any particular space, and hence is a necessary condition for the experience of any object in space. Every space we experience is experienced as bounded, and we can only experience a space as bounded if we are somehow implicitly aware that there is a space beyond the boundary that is not experienced. This un-experienced surrounding space is a condition for the possibility of the experienced space (being a space).

As a result of this every object of experience, being experienced as essentially in space and/or time, is experienced as pointing to something beyond itself that is not experienced. The awareness that there is something not experienced is a necessary condition of the experience of anything as in space or time. This awareness is an awareness that the immediate object of experience could not *be* (in space/time) if the non-experienced past time and surrounding space did not exist. The existence of past time and surrounding space, then, is a necessary condition for the existence of things in space/time. This is why every object of experience is experienced as conditioned. This, I believe, is the general structure of Kant's argument developed in the *Transcendental Deduction* and the *Analogies*. In the following two sub-sections I shall examine this argument in slightly more detail and attempt to justify my interpretation.

The *Transcendental Deduction*: The productive imagination is the name of our capacity for representing what is not present

Contained in every experience of objects, then, is a representation of something that is not experienced. In the *Transcendental Deduction* of the *Critique of Pure Reason*, Kant names the capacity to represent something that is not present the **imagination**, stating that, “*Imagination* is the faculty for representing an object **even without its presence** in intuition” (B151).¹⁴⁸ And he argues that what he calls the **productive imagination** is necessarily involved in all experience. There is much disagreement about how to read the *Deduction*. On my interpretation, the productive imagination is necessarily involved in all spatio-temporal experience because such experience essentially involves the awareness (or representation) that there is something that is not experienced (for example, past time and surrounding space). The ‘productive imagination’, then, is just Kant’s name for the capacity we have to, amongst other things, represent the past while we are experiencing something as happening now, and to represent the surrounding space while we are experiencing something as within our visual field. Such a representation of the past is involved in all experience of presence, for to experience a happening as the state of an object involves being conscious that the object exists in time and that there was some state the object was in prior to the state it is in now. An awareness that the object ‘was’ in some state or other is part of our awareness that what we are now experiencing ‘is’ the state of an object. Time is not a series of unconnected nows; for something to be present is for it to be essentially related to some past; for something to be past involves an essential relation to some present, for what it is to be past is to be the past of some

¹⁴⁸ “*Einbildungskraft* ist das Vermögen, einen Gegenstand auch ohne dessen Gegenwart in der Anschauung vorzustellen.”

present.¹⁴⁹ To experience something in time, then, is to experience something present but to be aware that it has a past. Of course, we are not (necessarily) aware *what* past it had, but, insofar as we are something as the present state of an object, we must be aware *that* it had a past. Now, Kant believes that he can explain how this capacity to be aware of the past (which is one aspect of the productive imagination) actually works, arguing that this capacity is judgmental. In other words, Kant believes that the productive imagination has a conceptual structure, governed by the categories. This is the point Kant is trying to make when he claims that the imagination “depends on understanding for the unity of its intellectual synthesis” (B164). Now, as the awareness that the object we are experiencing now *was*, is an awareness that the currently experienced state of the object has some **relation** to a past state that we are not currently experiencing, Kant concludes that this capacity must have something to do with our categories of relation, and he will argue that our awareness that what is being experienced has a past can be explained as an implicit *judgment* that what is being experienced is an *effect*. Kant believes, then, that everything that is experienced as present, then, must be experienced *as* an effect, and this is the conclusion he argues for in the *Second Analogy*. The relation between present and past, then, is to be understood in terms of the second category of relation.

The *Second Analogy* and temporal experience – to be experienced as happening in time is to be experienced as an effect

¹⁴⁹ In this saying this I am suggesting that Kant would agree with Sartre (2001) that “the three so-called “elements” of time, past present and future, should not be considered as a collection of “givens” for us to sum up – for example, as an infinite series of “nows” in which some are not yet, and others no longer – but rather as the structured moments of an original synthesis” (p.83).

One way of representing that which is not present is through analogy and thus the capacity to make analogies is to be understood as a particular instance of the faculty of imagination.¹⁵⁰ The *Transcendental Deduction*, then, provides a general account of the role of the (productive) imagination in experience whereas the *Analogies* should be read as examining in more detail the conceptual structure of the productive imagination. To illustrate this role of the imagination, and its conceptual structure, I will briefly examine Kant's *Second Analogy*, which involves an account of the role of the hypothetical form of judgment and of the corresponding categories of cause and effect in our experience of time. The purpose of this discussion is to further clarify why Kant maintains that everything we experience is experienced as conditioned.

In contrast to the Leibnizians, who argued that our experience of time is merely a confused representation of conceptual relations, Kant believes that there is something in the nature of our experience of time that is irreducibly non-conceptual. Thus, in the *Transcendental Aesthetic* Kant famously argues that time is “not a general concept, but a pure form of sensible intuition” (A31/B47). Time, then, is not reducible to logical or conceptual relations, but this is not to say that it does not have a conceptual structure. The logical structure of time is that of a series. Thus Kant maintains that,

¹⁵⁰ In his general introduction to the *Analogies*, Kant contrasts what ‘analogy’ signifies in mathematics with what it signifies in philosophy. He explains that, “in philosophy analogies signify something very different from what they represent in mathematics. In the latter they are formulas that assert identity of two relations of magnitude, and are always constitutive, so that if two members of the proportion are given the third is also thereby given, i.e., can be constructed. In philosophy, however, analogy is not the identity of two quantitative but of two qualitative relations, where from three given members I can give *a priori* only the relation to a fourth member but not this fourth member itself, though I have a rule for seeking it in experience and a mark for discovering it there” (A179-80/B222). An analogy then has the form: a is to b as c is to d. In mathematics if a, b and c are given, d is also, in a sense, given, in that we are able to construct d. In philosophical analogies if a, b and c are given we have a certain representation of ‘d’ (its relation to ‘c’), but ‘d’ itself is not given. We know that there must be a ‘d’ but we don’t know what it is.

Time is in itself a series. . . and hence in it, in regard to a given present the *antecedentia* are to be distinguished *a priori* as conditions (the past) from the *consequentia* (the future). (A413/B440)

Time, then, has a structure; the past is related to the future as antecedent to consequent. Now if time were merely a series of points we would not have to postulate a separate faculty of intuition, for the antecedent-consequent relation is a logical relation. Kant, then, believes that time is a series or an order and that the notion of a series (or of an order) can be **understood** in terms of the hypothetical (if. . . then) form of judgment, for he believes that the before-after relation can be reduced to the cause-effect (antecedent-consequent) relation. The notion of a series of points (that is, an order), then, can be fully grasped conceptually. Time, however, is **not** a series of points, but a series of times. These times have a relationship to one another that can (and must) be grasped conceptually by the understanding. However, although the *relation* between times is logical and can be fully grasped conceptually by the intellect, the experience of time cannot be reduced to the notion of an order of point-like time events, because what is ordered is not a series of (non-temporal) moments, but a series of times.¹⁵¹ The fact, however, that happenings in time are necessarily experienced as a series means that there is a conceptual relationship between these happenings. To perceive that something is happening now, is to be aware that something has happened before, which is to represent what is happening now as an effect. Kant makes this clear in the *Second Analogy*, where he argues that,

if I perceive that something happens, then the first thing **contained in this representation** is that something precedes. (A198/B244 – my emphasis)

¹⁵¹ One could make a similar point about space. Space does not consist of points (which Kant believes are limits of space rather than parts of it) but of spaces.

Here he makes it clear that it is not that I perceive something happening and then make a judgment that something precedes, but that the awareness that something precedes is contained in the perception of the happening. To recognize that something precedes what is happening now, however, is to recognize that what is happening now is part of a series of events. In other words, it is to recognize that

there is an order among our representations, in which the present one (in so far as it has come to be) **points to** some preceding state as a correlate, to be sure still undetermined, of this event which is given, which is, however, **determinately related** to the latter **as its consequence**, and necessarily connected with it in the temporal series. (A198/9 – My emphasis)

Here Kant makes it clear that any present happening points to and is determinately related to some preceding happening, and that the nature of this relation is that the present happening is (experienced as) a **consequence** of some past happening. The reason for this is because the way we experience something present as ‘pointing to’ some prior state, is to be aware of it as part of a series of events, and Kant believes that members of a series are ordered according to the ground/consequence relation. He makes this clear in his metaphysics lectures where he explains that,

The relation of consequence to ground [which is derived from the hypothetical form of judgment] is a relation of subordination; and things which stand in such relation constitute a series. Thus this relation of ground to consequence is a principle of the series, and is valid merely of the contingent. (*Metaphysic L2*, 28:551)

To experience a happening as present (now), then, is to experience it as happening in time. This, however, involves experiencing it as a part of a time series. We experience a

present representation (or happening) as part of a time order or time series, by experiencing it as pointing to some preceding happening. And the way we do this is to experience it *as* an effect. Everything experienced in time, then, is experienced as an effect. This is one reason why everything experienced in time is necessarily conditioned.

Kant sketches this argument, to be given in more detail in the *Second Analogy*, in section 26 of the second edition *Transcendental Deduction*. Here he explains that,

If. . . I perceive the freezing of water, I apprehend two states (of fluidity and solidity) as ones standing in a relation of time to each other. But in time. . . I represent necessary synthetic unity of the manifold, without which that relation could not be determinately given in an intuition (with regard to the temporal sequence.) But now this synthetic unity, as the *a priori* condition under which I combine the manifold of an **intuition in general**, if I abstract from the constant form of **my** inner intuition, time, is the category of cause, through which, if I apply it to my sensibility, I determine everything that happens in time in general as far as its relation is concerned. Thus the apprehension in such an occurrence, hence the occurrence itself, as far as possible perception is concerned, stands under the concept of the relation of effects and causes, and so in all other cases. (B163)

In the *Third Analogy*, Kant offers a similar account of our experience of objects in space. For to experience something as here, in front of us, involves an awareness of a surrounding space which is not here in front of us. I am always aware *that* my visual field is bounded, and *that* there is space that is not present to me beyond my visual field. I am aware *that* there is space beyond my visual field, but I am not aware of the space. The way a present space is experienced as pointing to some absent space, by means of which the experienced-space is experienced as surrounded by non-experienced space (and hence as bounded), is by being experienced as subject to the category of community. Once

again, this awareness is explained in terms of an explicit judgment, for the category of community is derived from the disjunctive form of judgment. Now, we always experience ‘this’ space as a **part** (of one single space),¹⁵² and Kant believes that we represent ‘this’ space as a part in the same way that we represent a part of a disjunctive judgment as part of a disjunction. A disjunctive judgment has the form: “x is A, or x is B, or x is C”, with the ‘or’ being understood as an exclusive or and the concepts A, B and C exhausting the possibilities. The implicit awareness that the space given in my visual field is only part of space, is, Kant believes, an implicit judgment that it is a part in the same way that “x is A” is part of the disjunctive judgment: “x is A or x is not-A”. One could say that every experience of a particular space, as a particular space, involves the implicit disjunctive judgment: “Space is either this particular space or it is not this particular space”. This judgment captures, and makes possible, the logical relations between spaces. Once again, however, our experience of space cannot be reduced to our understanding of the (logical) relations between spaces, for what are related are not logical points, but spaces. To be aware of this space as a space involves more than experiencing it as a part, for it is experienced as a space that is itself a part of space. In other words, although Kant believes that the part-whole relationship between spaces can be fully grasped conceptually in terms of the disjunctive form of judgment, space cannot be reduced to this relationship.

¹⁵² As Kant explains in the *Transcendental Aesthetic*, the parts of space “cannot as it were precede the single all-encompassing space as its components (from which its composition would be possible), but rather are only thought **in it**. It is essentially single; the manifold in it, thus also the general concept of spaces in general, rests merely on limitations” (A25/B39).

To conclude: Everything given in experience, then, is conditioned. Everything experienced as happening in time points to some (past) time that is the condition of the time experienced. Every experience of something in space involves an awareness that the space experienced is bounded, which involves an implicit awareness that there is some space beyond the space immediately experienced which is the condition of the space immediately experienced.¹⁵³ Experiencing the phenomenally given as conditioned involves the categories of relation, because to experience something as conditioned is to experience it as having a relation to something else. Such relations are not accidental. For something to be present it must have a past. If it there was no past there could be no present. The past, then, is a (logically) necessary condition of the present. “A single all encompassing space” is a (logically) necessary condition of this space being a space.

Ideas of pure reason, on the other hand, are ideas of things that are not conditioned, and so nothing that can be given in experience can ever be experienced as instantiating an idea.

¹⁵³ It is the condition of experiencing the experienced space as bounded.

Chapter Six

Theories of Interaction:

An Historical Overview

(6a) Introduction

The development of Kant's account of interaction has to be placed in its historical context. In eighteenth century German metaphysics textbooks, and in particular in Baumgarten's *Metaphysics* (the textbook Kant used for his metaphysics lectures), the question of interaction was dealt with under two headings: Psychology and Cosmology. The psychological question is about a particular type of interaction, namely that between mind and body, and is more familiar to contemporary philosophers. It developed in response to Cartesian dualism. Descartes argues that mind and body are two radically different types of substance, and the psychological question has to do with understanding how two such radically different types of substance can interact with each other. The cosmological question, in contrast, is more general, and asks how substances in general, even substances of the same type, can interact with each other. The cosmological question, although not as prominent today, was a major topic of debate in the eighteenth

century and remained of central importance to Kant throughout his career. For example, he argues in his *Inaugural Dissertation* of 1770 that,

The hinge, then, upon which the question about the principle of the form of the intelligible world turns is this: to explain how it is possible *that a plurality of substances should be in mutual interaction with each other, and in this way belong to the same whole which is called a world.* (2:407)

The critical Kant focuses almost exclusively on addressing the cosmological question, for by the 1780s he had come to see the psychological question as a non-question. By his critical period Kant had come to distinguish between inner and outer sense. Time is the form of inner sense, whereas both space and time are the form of outer sense.¹⁵⁴ For the mature Kant, the (phenomenal) self is just the object of inner sense, whereas body is the object of outer sense. Hence we have no reason to believe that there are two ontologically distinct types of substance whose interaction needs to be accounted for.¹⁵⁵ We merely have two types, or aspects, of sensibility. Kant's transcendental idealism, then, led him to the conclusion that the everyday distinction we make between the mental and the physical is not a distinction between two distinct ontological spheres and two corresponding types of substance, but instead merely a distinction between two forms of sensibility, with the physical being understood as the object of outer sense and the mental

¹⁵⁴ The claim that **both** space and time are the form of outer intuition is controversial, as Kant himself often seems to suggest that space alone is the form of outer sense. I believe that it makes more sense to regard both space and time as the form of outer sense, because the objects of outer sense are objects *in motion*. Kant himself explains the difference in these terms in the *Metaphysics of Morals*, where he argues that, "In theoretical Philosophy it is said that only objects of outer sense are in space, whereas objects of outer as well of inner sense are in time since the representations of both are still representations, and as such belong together to inner sense" (6:220).

¹⁵⁵ Thus Kant could argue in his lectures on metaphysics as early as the mid-1770s that, "the separation of the soul from the body consists in the *alteration of sensible intuition into spiritual intuition; and that is the other world*. The other world remains the same with respect to its objects; it is not different with respect to the substances, but is *intuited spiritually*" (*Metaphysik L2*, 28:298).

the object of inner sense. The question as to whether these two types of sensibility provide access to two different types of substance is a question that is in principle unanswerable. As Kant explains in his *Observations to the Paralogisms*:

Matter thus signifies not a species of substances quite different and heterogeneous from the object of inner sense (the soul), but rather only the heterogeneity of the appearances of substances (which in themselves are unknown to us), whose representations we call external in comparison with those that we ascribe to inner sense, even though they belong as much to the thinking subject as other thoughts do. . . Now the question is no longer about the community of the soul with other known but different substances outside us, but merely about conjunction of representations in inner sense with the modifications of our outer sensibility, and how these may be conjoined with one another according to constant laws, so that they are connected into one experience. (A385-6)

Thus, by the time of the first *Critique*, Kant had come to conceptualize the mind and body, not in terms of two distinct types of substance whose interaction then had to be explained, but in terms of two forms of sensibility. Through outer sense we are aware of bodies; through inner sense we are aware of our own ‘mental’ states. It would be a mistake, however, to conclude that these two different types of sensibility give us access to two ontologically distinct types of substance whose interaction then needs to be explained.

By the 1780s, then, Kant had concluded that there is ultimately no compelling reason to attempt to answer the psychological question. However, his attempts to answer the cosmological question and to explain the possibility of conceptualizing how “a plurality of substances should be in mutual interaction with each other” lies at the center of his philosophical development and had an enormous impact on the development of not only

his theoretical philosophy, but also his moral philosophy. Indeed, understanding Kant's final account of what is involved in the idea of 'real' interaction is the key to understanding much of his mature ethical writing, and in particular understanding the importance he places on the idea of autonomy. For he concludes that the only way we can conceive of a community of individuals really interacting is if we conceive of each individual member of the community as an autonomous agent. As a result, the decision to choose to interact really with others *is* the choice to be autonomous and vice versa.

(6b) The Three Standard Accounts of Interaction: Occasionalism, Pre-established Harmony, and Physical Influence.

By the time Kant began his philosophical career, there were three standard answers to the 'cosmological' question of interaction: pre-established harmony, occasionalism and physical influx or influence. This tri-partite division can be traced back to Leibniz. For example, in a letter to Basnage de Beauval (1696), Leibniz, in the context of a discussion about the psychological question of mind/body interaction, elucidates the three possible accounts of interaction by drawing an analogy with a pair of clocks.¹⁵⁶ Leibniz writes,

Consider two clocks or watches in perfect agreement. Now this can happen in *three ways*: the *first* is that of a natural influence . . . *The second way* to make two faulty clocks always agree would be to have them watched over by a competent workman, who would adjust them and get them to agree at every moment. *The third way* is to construct these two clocks from the start with so much skill and accuracy that one can be certain of their subsequent agreement. . . *The way of influence* is that of the common philosophy; . . . *The way of assistance* is that of the system of occasional causes. But, I hold, that is to appeal to a *Deus ex machina* in a natural and ordinary matter, where, according to reason, God should

¹⁵⁶Kant himself seems to have been aware of this passage and refers to this analogy while explaining Leibniz's position in his lectures on metaphysics. See *Metaphysic Mrongovius*, 29:866-7.

intervene only in the sense that he concurs with all other natural things. Thus there remains only my hypothesis, that is, *the way of pre-established harmony*.¹⁵⁷

Bayle in his *Historical and Critical Dictionary* popularized this tripartite taxonomy of theories of interaction. There are various ways of characterizing the difference between these three positions. The simplest is to explain it in terms of a finite substance's responsibility for (internal and external) change. Thus, the theory of (*physical*) *influence* asserts that individuals can cause changes both in themselves and in others – that is, they can cause both internal and external change. The theory of *occasionalism* denies that finite substances are the cause of change either in themselves or in others.¹⁵⁸ The theory of *pre-established harmony* asserts that finite substances are the cause of changes in themselves, but not in others. In addition, it should be noted that in the early eighteenth century the dominant account of 'physical influx' involved (as the name suggests) the idea of the accidents of one substance 'flowing into' another substance.¹⁵⁹ Following Kant I will refer to this position as the theory of *crude physical influence*. Kant, although

¹⁵⁷ Leibniz (1989), pp.147-8. It should be noted that in this letter Leibniz is discussing the psychological question of mind-body interaction and not the cosmological question of interaction in general. Kant, in his lectures, however, appeals to this analogy during his discussion of the cosmological question. And Leibniz himself ultimately regards the psychological question as a special instance of the cosmological question.

¹⁵⁸ It is not clear if this reading is fair to Malebranche. In Germany in the eighteenth century it seems that the standard reading of Malebranche was that he denied all activity to finite substances. Malebranche himself seems to have believed that finite individuals do have wills in that they possess the capacity to assert or deny. I think most of his readers in the post-Leibnizian tradition took this to mean that Malebranche denied that finite substances were *active in any meaningful sense*.

¹⁵⁹ Leibniz traces the crude theory of physical influence back to Suárez. Duarte (2001), however, argues that this attribution does not stand up to close scrutiny. Suárez does explain action in terms of influence rather than dependence, however 'influre' is a transitive verb the direct object of which is 'esse' (being), with the patient being acted upon taken as the indirect object of this verb. Suárez does not, then, suggest that in interaction something (say an accident) is transferred from the agent to the patient. What, Suárez actually meant by the notion of the agent "flowing being into" the patient, is unclear. See Disputation 12 of his *Metaphysical Disputations, De Causis Entis in Communi*, masterfully translated by Shane Duarte (unpublished).

he rejects crude physical influx, can be understood as advocating a version of physical influx.

Watkins (1995b) explains the trichotomy in the following terms:

Physical influx asserts intersubstantial causation amongst finite substances. For instance, when I appear to kick a ball, I really am the cause of the ball's motion. Pre-established Harmony denies intersubstantial causation, but affirms intrasubstantial causation. According to Pre-established Harmony, then, I am not the cause of the ball's motion, but rather the ball is simply causing itself to move . . . Occasionalism, like Pre-established Harmony, denies intersubstantial causation, but, unlike Pre-established Harmony, it denies intrasubstantial causation as well. Occasionalism typically asserts that God alone, that is, an infinite substance, is the cause of all changes, and thus of the ball's motion. (p.296)

Watkins suggests that one advantage of these definitions is that the theories form an exhaustive disjunction, and he rightly points out that this is how the theories were generally understood in eighteenth century Germany. At this stage in my investigation I will suggest that these definitions provide a good starting-point, and that at least at the beginning of his career, this seems to be how Kant understood the distinction between the three positions.

By the mid eighteenth century, at least in the German-speaking world, this exhaustive tripartite taxonomy was pretty much taken for granted. Thus, as Watkins (1995a) notes, by 1723 Bilfinger could claim that occasionalism, pre-established harmony and physical influence were the *only* three possible theories of interaction.¹⁶⁰ And it seems fair to say that in the German milieu in which Kant developed philosophically, only two of these

¹⁶⁰ See Watkins (1995a), p.285.

answers were regarded as serious contenders: pre-established harmony and physical influence. Following Leibniz, Wolff and his school tended to be defenders of pre-established harmony. On the other hand a number of important philosophers such as Gottsched, Crusius, and Knutzen (who was one of Kant's teachers in Königsberg) had written tracts advocating physical influx. Occasionalism was no longer taken seriously.¹⁶¹ Kant himself rejects occasionalism without much discussion and I suggest that this dismissal was motivated by the fact that he regarded individual substances as essentially active (or by the critical period as 'agents') which led him to see occasionalism, understood as a doctrine which denies the real agency of finite individuals, as essentially a denial of the possibility of finite individuality and hence as akin to Spinozism.

(6c) Kant's Rejection of Pre-established Harmony

Kant's rejection of pre-established harmony can be traced back to his first published work, the *True Estimation of Living Forces* of 1746. In this work he speaks approvingly of the "triumph" of physical influx over pre-established harmony (1:20), and Kant remained a committed opponent to pre-established harmony throughout his career. Despite his unwavering rejection of pre-established harmony, he consistently agreed with Leibniz that individuals (or monads) are 'windowless'. And as a result of this he consistently rejected crude physical influx and the thesis that in cases of interaction

¹⁶¹ The reason for this seems to have been because following Leibniz most German metaphysicians of the period conceived of substances, even finite substances, as essentially active. Insofar as occasionalism was taken to deny the activity of finite substances it was taken to imply a contradiction.

something flows from one individual ('through a window') into another individual. He was, and remained, committed to the 'windowless' nature of individual substances because he was committed to what I shall call the **Principle of Active Inherence**. According to this principle, an accident, determination¹⁶² or state can only be an accident, determination or state of a particular individual if the individual itself is the 'ground' or 'cause' of the state. Kant, however, did not believe that accepting this principle implied rejecting the possibility of real interaction.

Kant offers an interesting argument against pre-established harmony in the *New Elucidation* of 1755. In Section Three of this work he discusses the 'principle of succession' and attempts to justify the proposition that, "no change can happen to substances except in so far as they are connected with other substances" (1:410). The structure of his argument is simple. The defender of pre-established harmony believes that finite individual substances are isolated from each other but that they experience changes of states. Kant attempts to show that it is impossible to explain how changes of states are possible in an isolated substance. It is not clear, however, whether he offers three formulations of a single argument or three distinct arguments. The initial (formulation of) the argument is as follows:

[1] The inner determinations, which already belong to the substance, are posited in virtue of inner grounds which exclude the opposite. Accordingly, [2] if you want another determination to follow, you must also posit another ground. But since the opposite of this ground is internal to the substance, and since, in virtue of what we have presupposed, [3] no external ground is added to it, it is patently obvious that [4] the new

¹⁶² The term 'determination' is Kant's own. An accident, property or state of a substance would all fall under the term 'determination'.

determination cannot be introduced into the being. (1:140 - numbers added)

The first premise is what I have named the Principle of Active Inherence, and Kant remained committed to it throughout his career. The crucial step in this argument is, Kant believes, the second one, which he rejects. For, he will argue that for a new determination to occur does not necessarily require the existence of a new ground, but it might be the consequence of a pre-existing ground that has not been able to actualize itself because of some external resistance. The easiest way of explaining this is in terms of an example. For the sake of the argument I will assume that motion is an inner determination. Kant would agree that the motion of an individual substance must be due to some inner ground, and calls the ground of a determination a force. Therefore, Kant would accept that there must be some force internal to the substance that is responsible for its motion. A *change* of motion, however, does not require the existence of some new internal ground or force. For the change of motion may be due to a withdrawal of resistance, which allows some already existing, but ineffective, internal ground to become effective.

Kant, then, in advocating the principle of active inherence, agrees with the Leibnizians that individual substances must be thought of as windowless. He agrees that the determinations (say, the motion) of a substance must be due to inner grounds. His position, however, diverges from Leibniz' in that he believes that this does not imply that a *change* of determinations must also be due to grounds internal to an individual

substance. Indeed, he believes that the defender of pre-established harmony has no way of explaining how any change of states in a substance is possible.

The defender of pre-established harmony, however, could reply that she doesn't believe that that it is the 'determinations' or states of a substance that depend upon inner grounds, but, instead, that it is the *change* of states or determinations that must be the results of inner grounds or forces. If this is the case, then the defender of pre-established harmony has no problem explaining the possibility of change. And, indeed, this is precisely what Leibniz and Wolff argued. And, moreover, Kant himself recognizes this. Thus, he argues that although he is familiar with the arguments of the Leibnizians he is "convinced of their sterility", because,

once they have constructed an arbitrary definition of force so that it means that which contains the ground *of changes*, when one ought to declare that it contains the ground of *determinations*, they were bound to fall headlong into error. (1:411)

Thus Kant recognizes that the Leibnizians define a substance's force as the ground of its changes and not merely as the ground of its determinations or states. If we are willing to accept this conception of force, however, it would seem that Kant's argument in the *New Elucidation* misses the mark, for if we understand the 'force' internal to a substance as the ground of its changes, we do not face any problem in explaining how change is possible in an isolated substance. This problem only arises if we accept Kant's definition of (internal) force as the ground of substance's determinations and not as the ground of its changes. And, unfortunately, Kant does not offer much in the way of arguments as to why we should regard the Wolffian definition of force as the ground of changes as any

more arbitrary than his own definition of force as the ground of a substance's determinations.

Although Kant's arguments in the *New Elucidation* are not compelling, this discussion should at least establish that in 1755 Kant himself believed that he had a decisive argument against pre-established harmony. Indeed, he felt justified in claiming that this "proof utterly overthrows the Leibnizian pre-established harmony" (1:412). Kant remained remarkably consistent on this issue for the rest of his life. For, as we shall see, he consistently advocates the position that an individual substance must be the ground of its own determinations (or states), but this does not imply that it must be the ground of its changes of determinations (or states.)

It is clear, then, that from the start of his career Kant believed that he had good reasons for rejecting pre-established harmony. It was only late in his career, however, that he developed a model of real interaction that was able to explain how the position he advocated can be meaningfully distinguished from that of pre-established harmony.

(6d) Kant's Attitude Towards Physical Influx

As we have seen, Kant explicitly rejects pre-established harmony and occasionalism. His attitude towards physical influence is less clear. In the *New Elucidation* he suggests that his own account is "certainly somewhat superior to the popular system of *physical influence*" (1:416). Ameriks (1992) takes this to mean that Kant sees his position as "a

fourth alternative” (p.262). Watkins (1995a), on the other hand, suggests that even in this paper Kant should be understood as proposing “a more sophisticated version of Physical Influx” (p.292). On Watkins’s interpretation, then, Kant is to be regarded merely as rejecting the ‘popular’ exposition of physical influx and not the theory of physical influx itself. In his later works, as both Watkins and Ameriks point out, Kant did often refer to his own account of interaction as a version of physical influence. Whatever he calls his position, he does offer a much more sophisticated theory of interaction than any of the three traditional alternatives, and as a result it is potentially misleading to refer to the position he advocates as physical influence. Although I agree with Watkins that, even in his earliest works, Kant understood himself to be providing a defense of physical influence broadly understood, I will generally refer to the position Kant advocates as a theory of *real interaction*. This is a phrase Kant often uses and it etymologically presents his mature position far more accurately than the expression ‘physical influence’. For the theory Kant wishes to defend is a theory of inter-action and not a theory of in-flowing. And the term ‘real’ in this context is used in a technical sense.

Although it is potentially misleading to regard Kant as a defender of physical influence, Kant himself often calls the position he is advocating physical influence. The reason for this is that the mature Kant had no problem with adopting the terminology of his day for polemical purposes. In claiming to be a defender of physical influx, however, he is merely attempting to indicate to his audience that he rejects pre-established harmony and occasionalism. He recognizes, however, that his use of this expression is potentially misleading, for his account is neither ‘physical’ nor does it involve any ‘in-flowing’. On

the other hand, understanding Kant as offering a defense of ‘physical influx’ broadly understood does help us to understand where Kant located himself in the contemporary controversy. Although he does at times label the position he advocates ‘physical influence’, he is clearly unhappy with this choice of words. He makes this clear on a number of occasions. For example, he points out in his lectures on metaphysics that, “influence <*influxus*> is an unfitting expression, it implies that the accident migrated out of a substance.” (*Metaphysik Mrongovius*, 29:823) This unfittingness, as we have seen, does not stop Kant from at times using this terminology. The mature Kant, then, uses the phrases ‘physical influence’ and ‘real interaction’ interchangeably to refer to the same doctrine.¹⁶³

Like Leibniz, Kant can be thought of as developing a ‘monadology’ or an account of what it is to be an individual. However, whereas Leibniz thought that his monadology offered an account of the way the world really is, the critical Kant is merely concerned with examining what is conceptually involved in the pure idea of an individual.¹⁶⁴ Such an account is not, however, a pointless exercise akin to building castles in the air because

¹⁶³ I should point out at this point that Kant’s use of *ideal*, as an adjective and contrasted to ‘real’ should be clearly distinguished from his use of this word as a noun, as in when he talks of ‘*an ideal*’. Although Kant uses the same word in both contexts the concepts are totally distinct and have nothing to do with each other. ‘Ideal’ in the phrase ‘Kant’s moral ideal’ has nothing in common with the ‘ideal’ in the phrase ‘ideal interaction’.

¹⁶⁴ The pre-critical Kant (before 1770) thought that his account of individuality (and hence his account of interaction between individuals) was an account of the way things are. The critical Kant, on the other hand, thought that the idea of an individual is an idea of pure reason and as such is not the sort of thing that can be an object of experience. That is, for the critical Kant, there are no real individuals to be found in the spatio-temporal phenomenal world. As he explains in the *Critique of Practical Reason*: “[I]f this ideality of time and space is not adopted, nothing remains but Spinozism, in which space and time are essential determinations of the original being itself, while the things dependent upon it (ourselves, therefore, included) are not [individual] substances but merely accidents inhering in it” (5:101-2 – my addition in square brackets). This passage clearly indicates that the critical Kant believes that there are no real individual substances to be found in the spatio-temporal phenomenal world.

the idea of what it is to be an individual plays an essential role in Kant's ethics. While an individual substance can never be an object of possible experience, the idea of an individual monad (in interaction with other individuals) does offer us an 'image' of what we can (and should) choose to be. Thus, it is important to clarify what is involved in this idea. Throughout his career Kant agreed with Leibniz that individuals are 'windowless'. That is, he agreed that the notion of a 'determination' flowing from one individual into another is incoherent.

One problem with the crude theory of physical influence is that it suggests that determinations (or in the traditional vocabulary 'accidents') are the sort of things that can 'float around' and exist independently of individual substances. Both Leibniz and Kant have a problem with this notion. A defender of crude physical influence, however, might just bite the bullet and accept the coherency of the idea of accidents existing independently of substances. A more serious problem, however, with the doctrine of crude physical influx is that it is unclear in what sense a determination can really be thought of as being a determination of either the agent or the patient. For if accidents or determinations are the sort of things that can detach themselves from individuals and float from one individual into another then we need some account of the way in which an accident can really 'stick to' or really belong to a particular individual in a way strong enough to make the accident an accident *of* that individual. If accidents are the sort of things that can be detached from an individual, we need to give some account of the real unity of accident and individual. Following Leibniz, Kant rejects the crude theory of physical influence, for he believes that the only way an accident can truly belong to (or

be unified with) a substance is if the substance is the (active) ground of the accident.¹⁶⁵ Kant's main criticism, then, is that the proponent of crude physical influx has no way of accounting for inherence.

Although Kant agrees with Leibniz that individual substances must be 'windowless' he does not believe that this implies that we must conclude that we cannot conceive of 'real' interaction between individuals. It was not until the 1770s or early 80s, however, that Kant was in a position to explain what he means by 'real' interaction, and so until this point in his career he is not really able to explain how his account can be distinguished from pre-established harmony.

The mature Kant reached the conclusion that individuals can only interact if they are members of a community, and that real interaction can be distinguished from ideal interaction in that in the case of real interaction the individuals constitute a *real*, as opposed to an *ideal*, community. The reasons he reached this conclusion will be the topic of the following two chapters. According to this position, the defender of pre-established harmony can be thought of as claiming that the community is ideal in the sense that its unity exists merely in the mind of God, whereas the defender of real interaction believes that its unity is real, constituting some fact about the world. Kant himself often presents the distinction between a 'real' and 'ideal' community (or 'totality') in these mentalistic terms and I suspect that this is maybe how he first started to conceive of this distinction. It is, however, not a very informative way to explain the distinction, for the defender of

¹⁶⁵ An interesting point to note here is that this seems to imply that the first category of relation (substance/accident/inherence) involves the second (ground/consequence).

pre-established harmony can always ask what is meant by the claim that in a 'real' community its unity constitutes a fact about the world. The defender of pre-established harmony could, quite justifiably, argue that the distinction between him and his opponent is merely verbal, for his opponent cannot explain what is meant by the term 'real'. Early in his career this criticism can justifiably be made against Kant, and the distinction between his position and that of a defender of pre-established harmony is purely verbal. Ultimately, however, I believe Kant worked out how to explain the difference between a 'real' and 'ideal' community in non-mentalistic terms. A 'real' community is to be distinguished from an 'ideal' community in that the unity of a real community is *intrinsic* to the community. What is meant by this is that the members of a real community must, by definition, themselves be the source of the relations (or laws) that provide the community with its unity. It is only once Kant has worked out how to explain the distinction between a real and ideal community in these terms that he is able to distinguish meaningfully his position on interaction from pre-established harmony.

To conclude: Kant accepts the Leibnizian claim that (individual) substances must be thought of as 'windowless'. He is committed to this because he is committed to what I have called the Principle of Active Inherence. According to this principle a 'determination' can only be the determination of a particular substance if the substance itself is the (active) ground of the determination. If we accept the Principle of Active Inherence we must reject the crude theory of physical influx. According to this theory when individual A (the agent) affects individual B (the patient) a determination (say an 'accident') "flows" from the agent into the patient, and so the patient is not the ground of

the determination. The problem with the crude theory of physical influx, then, is that it is difficult to understand in what sense the determination can really be a determination of the patient. Kant believes, however, that accepting the Principle of Active Inherence, although it rules out the intelligibility of crude physical influx, does not commit him to something like pre-established harmony.

To understand how Kant can accept this principle and still meaningfully defend some version of real interaction, it is necessary to examine his account of action in more detail. Understanding his account of action will also help us understand why interaction presupposes community, in the sense that individuals in interaction must be thought of as constituting a community.

Chapter Seven

Kant's Theory of Action

In this chapter, I will attempt to explain and justify the claim that interaction is only possible between individuals that constitute a community. In the following chapter I shall examine Kant's account of the idea of community in more detail, in particular examining its relationship to the category of community and the disjunctive form of judgment.

(7a) Introduction

As we saw in the previous chapter, Kant rejects the crude theory of physical influx, but does not believe that this means he must commit himself to either occasionalism or pre-established harmony. According to the crude theory of physical influence a substance *a* is the cause of a 'determination' or accident in substance *b* if the accident 'flows' from substance *a* into substance *b*. Such a theory might sound like an implausible straw man, but it reflects how we often talk about interaction. For example, we talk of motion being 'transferred' from one body to another. Similarly we talk of a property right being

transferred from one agent to another.¹⁶⁶ Also, many people conceptualize what is involved in a good conversation or a good teacher-pupil relationship in these terms. I have an idea in my head and I wish to ‘convey’ it to my conversational partner or student. On this model, successful communication (or teaching) involves the transferal of ideas from one individual to another. The idea somehow flows from my mind ‘through a window’ into your mind. Rejecting the coherency of physical influx and the idea that individuals have ‘windows’ (through which, say, motion, property rights, or ideas could flow in) implies that we need to reconceptualize these cases of ‘action’.

Kant rejects such a model of interaction because it is incompatible with his understanding of what it is for an accident to inhere in a substance. Given his rejection of crude physical influence, however, Kant needs to explain in what sense one substance can really act upon another.

The idea of interaction clearly involves the idea of reciprocal causation. Individuals *a* and *b* interact if *a* has a causal influence on *b* and *b* has a causal influence on *a*. It cannot, however, be reduced to the notion of reciprocal causation, because interaction essentially involves action, and Kant believes that only way of making sense of the action of one individual upon another is in terms of the withdrawal of resistance. On this account an individual *a* can only act upon, and be the cause of a change in, another individual *b* if the change of state in individual *b* is the result of individual *a* ‘withdrawing resistance’. This

¹⁶⁶ In Chapter 8 I will argue that a rejection of the coherency of crude physical influx implies that we are left with the task of cashing out what is involve in the ‘transferal’ of property rights in some other terms. I will argue that this is precisely what Kant is attempting to do in the *Doctrine of Right* of the *Metaphysics of Morals*.

account of interaction, however, presupposes that for interaction to be possible, the individuals interacting must somehow already be limiting or resisting one another. Kant believes that the concept of ‘resistance’ is a pure *a priori* concept, being what he calls a predicable of the category of community. So insofar as action presupposes action and action can only be understood as the withdrawal of resistance, interaction presupposes community. And this implies that interaction is only possible between individuals that already constitute a community.¹⁶⁷ This, however, leaves Kant’s account with a problem.

To understand this problem it is helpful to examine the problem Leibniz had explaining the possibility of composition. Leibniz believes that individual substances (monads) are isolated and cannot really interact. As a result of this he has a problem in explaining how a world of individuals is possible, for a world is a composite, and if individuals are really isolated it is difficult to explain their composition. Kant agrees with Leibniz that our idea of a world is the idea of a composite and he often argues that it is the fact that individuals interact that makes it possible for them to be members of a unified community. If this were the only account he could give of community membership then his argument would be viciously circular. For we cannot claim that individuals are members of a world by virtue of the fact that they interact if we believe, as Kant does, that individuals can only interact if they are members of a world. Given Kant’s account of action, the conceivability of such a community must be the basis for the conceivability of real interaction, and not *vice versa*. Fortunately, although Kant frequently claims that it is the real interaction between individuals that is the basis for their belonging to a unified

¹⁶⁷ This is because the concept of resistance is what Kant calls a ‘predicable’ of the category of community.

community or world, his analysis of the idea of community, which will be examined in the following chapter, explains how we can conceive of a multitude of individuals belonging to a unified community, without having to appeal to the fact that they are interacting. On this account, Individuals can form a community by willing a common set of juridical laws, and these laws are the basis of real, intelligible interaction for such laws (once they have been willed) are the source of real resistance between individuals.

(7b) A Problem

There are times when Kant seems to conceive of interaction merely as reciprocal causation. For example, in his metaphysic lectures he explains that,

The relation of community is different from that of the consequence to the ground in this, that the cause and effect are reciprocal here, i.e., there is something in the effect <causato> which is the ground of the cause <causa> and something in the cause <causa> which is ground of the effect <causato> = each concurrently. (*Metaphysik Vigilantius*, 29:986)

It is clear, however, that the concept of community involves more than the idea of reciprocal causation, for interaction involves not merely causation, but *action*.

The problem with conceptualizing (reciprocal) action is fairly simple. Following Leibniz, Kant thinks that the idea of an individual (substance) is the idea of something essentially active. There seems to be, however, a problem in conceiving of two essentially active beings mutually acting upon one another, and thus there is a problem in conceiving of a community of essentially active individuals. For in a community of

individuals in interaction individuals will not only act, but also be acted upon. The problem then is to give an account of how we can conceive of an essentially active substance as suffering or passive. Any account of interaction, then, must be able to explain how an agent can be a patient. Kant himself explicitly poses this question in his lectures on metaphysics. He explains that “that substance suffers (passive) whose accidents inhere through another power.” He then asks,

How is this passion possible, since it was said earlier that it [i.e. the passive/suffering substance] is active insofar as its accidents inhere. (*Metaphysik Mrongovius*, 29:823)

The problem, then, is not merely that Kant conceives of individual substances as essentially active, but that following Leibniz he is committed to a particular conception of inherence. Namely, he is committed to the view that an accident (or what Kant refers to as a ‘determination’) can only truly inhere in or belong to a substance if the substance is the active cause or ground of the accident. A determination, then, is only the determination of a particular individual if the individual is somehow the ‘ground’ of the determination. I have named this doctrine the Principle of Active Inherence.¹⁶⁸ If we accept the principle of active inherence, though, it is not clear how one individual can ever be the cause of any change in another individual. If a determination can only be a determination of individual *b* if *b* is the active ground or cause of the determination, how can another substance ever be the cause of a change in *b*? Leibniz’s solution was to admit defeat and conclude that one substance cannot be the cause of a change in another.

¹⁶⁸ I suspect that Kant and Leibniz are drawn to such a doctrine because of their worries about unity. An individual must be essentially unified. Given this fact there must be something that accounts for the unity of a substance and its ‘determinations’. The principle of active inherence provides such an account.

(7c) Kant's Solution

Kant's solution to this problem will be to claim that we can understand the idea of an individual being acted upon, without appealing to the untenable notion of accidents flowing into the individual, in terms of the agent "determining the active power of the substance being acted upon" (29:823). This account of action does not violate the Principle of Active Inherence, because the patient's determination inheres in the patient because it is a result of the patient's power. This power, however, has been determined by the agent. It is not clear, however, what we should make of this notion of the agent "determining the power" of the patient.

The model Kant introduces to clarify the notion of one individual determining the power of another is that of the withdrawal of resistance. Individuals, on this model, already resist one another. And one individual substance (the agent) is the 'cause' of a change in another individual substance (the patient) if the change in the patient is the result of the agent withdrawing its resistance. The patient remains, however, essentially active, for the determination is the result of its power. Thus each individual is essentially active in that everything that happens to a particular individual (everything a particular individual suffers) is the result of its own power or potentiality. But much of what we do (everything that 'happens' to us) only occurs when other individuals remove impediments. One could say that I am, in a certain sense, the ground of the possibility of all of my determinations, but am not the sole ground of all of their actuality.

Thus Kant explains in his lectures on metaphysics from 1782-3,

We can never be merely passive, but rather every passion is at the same time action. . . Every substance is self-active, otherwise it could not be substance;. . . The substance being acted upon <*substantia patiens*> is acting in itself <*eoipso agens*>, for **the accident would not inhere if the substance had no power through which it inhered in it**, hence it also acts; influence <*influxus*> is therefore an unfitting expression, as it implies that the accident migrated out of a substance. What then is genuine passivity? The acting substance <*substantia agens*> determines the power of the substance being acted upon <*substantiae patientis*> in order to produce this accident, therefore all passivity <*passio*> is nothing more than the determination of the power of the suffering substance by an outer power. (*Metaphysik Mrongovius*, 29:823 – My emphasis.)

Here Kant spells out his commitment to what I have called the principle of active inherence: “an accident would not inhere if the substance had no power through which it inhered in it.” And he believes that commitment to this principle rules out the possibility of crude physical influence. However, he does not believe that it rules out any commitment to real interaction, for it still allows for some account of passivity. An individual can be a patient, that is, it can be acted upon, if another individual “determines the power of the substance being acted upon in order to produce the accident”.

He continues his discussion by giving a more detailed account of how one individual can determine the power of another. This discussion is very similar to the account he gives in the course of his discussion of the question “what is acting?”, in his lectures on metaphysics of 1790-1. And as this later account is more compact I will move to a discussion of these later passages. Here Kant makes it clear that one individual ‘determines the power’ of another when it removes an impediment which allows a ‘dead’ power to become a living power. In this discussion, the language of which seems to have

been influenced by Leibniz's *Specimen of Dynamics*, he clarifies the distinction between a 'faculty' and a 'power' and explains the difference between a 'living' and a 'dead' power. He explains:

With a faculty we imagine only the possibility of power. Between faculty and power lies the concept of endeavor <*conatus*; *Bestrebung*>. When the determining ground for an effect is internally sufficient, then it is a dead power. But when it is internally and externally sufficient, then it is a living power. **Power which is merely internally sufficient, without being able to produce the effect, is always opposed to an opposing power which hinders its effect, an impediment <*impedimentum*>. Thus as soon as the impediment <*impedimentum*> is removed, the dead power becomes living.** (*Metaphysik L2*, 28:565)

Here Kant distinguishes between the idea of a 'faculty' and the idea of a power. A faculty is a mere capacity whereas a power is already a striving or endeavor. 'Conatus' is a term that Kant has borrowed from the calculus. Imagine a ball at rest. It has the faculty or capacity to move in a straight line. Now, imagine a ball being swung attached to a rope and swinging round a fixed point (or better, imagine the moon attracted to the earth by the force of gravity and circling it). At each particular moment the ball 'wants' to move in a straight line, at a tangent to the circle it is describing. This is what Leibniz termed 'conatus' and what Kant refers to in German as 'endeavor' (*Bestrebung*). Thus, although the ball is actually moving in a circle, at any particular moment it is 'endeavoring' to move in a straight line along the tangent. At any particular moment it would move along the tangent if all external forces were removed. Kant calls this 'endeavor' to move along the tangent a 'power'; it is more than what Kant calls a capacity or faculty, for even an object at rest has the capacity to move along a straight line. We can, however, distinguish between a 'dead' power and a 'living' power. The power of the ball (to move along the

tangent) will remain a ‘dead power’ unless the rope is cut. If the rope is cut, the impediment is removed, and the ball will move off along the tangent. Upon the cutting of the rope the dead power becomes a living power. The cutting of the rope ‘causes’ the ball to fly off in a straight line – but this cutting merely allows for the actualization of the ball’s dead power. So the motion of the ball along the tangent really is the ball’s motion.

Thus, a static point has the *capacity* or *faculty* to move along a straight line. If it is moving in a circle around a center of gravity at every moment it is ‘striving’ to move along the tangent. In such situations, at each moment it has a *dead power* to move along the tangent. If the force of gravity is removed it will move along the tangent along a straight line. In moving along a straight line it is exercising a *living power*. Although the movement along the straight line is due to its own power, the removal of the force of gravity is the *cause* of its motion in a straight line.

Kant suggests that all interaction between substances can be understood as analogous to this. Of course, the example of the ball I have just given is a spatio-temporal one. The analogy, however, can be stripped of its spatio-temporal elements and be applied to intelligible individuals, that is to individuals insofar as they are thought as not subject to the spatio-temporal conditions of our form of intuition. Thus, one individual (the agent) acts upon another (the patient) when the agent is the ‘cause’ of a dead power in the patient becoming a living power. The agent can only be the cause of such a change by removing an impediment, which allows a dead power to change into an active power. In so doing, the agent causes a change of determination in the patient. The (new)

determination, however, belongs to the patient because it is a result of the patient's dead power or 'endeavor'. The agent only 'causes' the determination by removing some impediment. This explains how an essentially active individual can be acted upon and hence how reciprocal causation is possible, and why the notion of interaction involves the concept of community, for only individuals that are (already) members of a community can impede one another.¹⁶⁹

As I have argued that, for Kant, the only type of individuals that can really interact are moral (autonomous) agents, it is worth examining how this model of action can be applied to moral agency. In particular it allows us to draw the distinction between someone's being the cause of an action and someone's being responsible for an action. Suppose that my next door neighbor is a rich art dealer. Every evening I walk past his house on the way back from the pub. I see a Picasso above his mantelpiece, and want it. Occasionally I try his front door, but it is locked. One day the art dealer forgets to lock the door, or perhaps he decides that he should be more trusting and decides to leave the door unlocked. I steal the painting. If the art dealer had not left the door unlocked, I would not have stolen the painting. In such a case I think that we would be tempted to say that the fact that the art dealer unlocked his door was, at least in one sense, the cause of my stealing the painting. At the same time we do not want to hold the art dealer

¹⁶⁹ It is interesting to apply this analogy and terminology to Kant's ethical writings. We could say that consciousness of the moral law shows that we have the capacity to be autonomous, that we possess the **faculty** of practical reason. To be moral is to **endeavor** to follow the moral law. The most we can achieve alone, however, is a dead power. What we must hope for is that the dead power the endeavor to follow the law creates is allowed to become a **living power**. A moral world is a world in which each individual follows his or her own course without impeding others or being impeded. In a kingdom of ends there would, so to speak, be no dead power. An individual is happy when she realizes her goals – in this language we could say that she would be totally happy only if all of her power were living power. This is only possible, however, in a world in which all individual submit themselves to the same laws.

responsible for my action. His action may have caused the theft, but he is not responsible for the theft. I am responsible for stealing the painting, Kant would argue, because of my character; perhaps one of my maxims is to “take anything I desire if I believe I will not get in to trouble”. The action of the art dealer does not cause a change in my character. I will only take the painting if I am already pre-disposed to behave in such a way if the opportunity arises. The decision of the art dealer, however, is the cause of some change in me. I have gone from wanting to possess the painting to actually possessing it, from being a ‘law abiding citizen’ to being a thief. We could say that the act of the art dealer in unlocking the door is the cause of my change of state (say, from being a non-thief to being a thief) but he is not responsible for my being a thief. The action of others, then, allows my character to reveal itself in certain ways (both to the world at large and to myself) but what is being revealed is *my* character or disposition.

(7d) Leibniz’ Problem with Explaining the Possibility of Composite Substances

Unlike Leibniz, then, Kant believes that he can explain how individuals can be thought of as interacting without violating the Principle of Active Inherence. He also believes that the fact that we can conceive of individuals as interaction allows us to understand how individuals can be members of a unified world or community. For our idea of a world is the idea of a composite whole consisting of individuals, and Kant maintains that it is the fact that individuals are thought of as interacting that allows us to conceive of them as members of a whole. Thus, in his lectures on metaphysics, in the course of discussing our idea of an intelligible world, he explains that,

The world is composite because it has a multitude of substances, and whole <totum> because all of these stand in interaction <commercio>. (*Metaphysik Mrongovius*, 29:852)

Here, then, Kant suggests that it is the fact that individuals interact that explains how they can be members of a whole. Similarly in his lectures from the early 1790s he argues that,

The aggregation of the substances in which there is no community still does not constitute a world. Reciprocal determination, the form of the world as a composite <compositum> rests upon interaction <commercio>. (*Metaphysik L2*, 28:196)

Similar passages are not hard to find.¹⁷⁰ Unfortunately, Kant's position is inconsistent. Kant's argument here is circular, for as we have seen his account of action presupposes that individuals in interaction must be members of a whole. If there was not some connection between individuals, how could individuals be impeding one another? So the fact that they are thought of as interacting cannot be used to explain how we can think of them as members of a composite whole. In the following chapter, I will argue that Kant is able to explain how individuals can be thought of as members of a whole (as members of a community) without appealing to their interaction. Before moving on to this topic, however, I will spend the remainder of this chapter examining why a philosopher working in the Leibnizian tradition has problems with explaining the possibility of composition and why Kant's own account of action means that the fact that individuals interact cannot account for the unity of the whole they are part of, for interaction between individuals presupposes that they are members of a unified whole, and so cannot be the basis of the unity of the whole. In the following chapter I will explain how Kant believes

¹⁷⁰ This type of claim can be traced back to the Inaugural Dissertation, where Kant argues that the problem he wishes to examine is, "How it is possible that a plurality of substances should be in mutual interaction with each other, and **in this way** belong to the same whole which is called a world" (2:407 – my emphasis).

it is possible to conceive of such a unified composite whole of individuals through the category of community. And it is only because we can think of individuals as members of a (real) community, conceived of in these terms, that we can think of them as interacting and not *vice versa*.

Leibniz himself recognized that his commitment to pre-established harmony made it very difficult, if not impossible, for him to explain mind/body unity, and the problems he finally recognized with his account of mind/body unity are equally telling against the unity of a world of Leibnizian monads. For if pre-established harmony cannot account for the union of mind and body, it is equally incapable for accounting for the unity of a 'world' of monads. And Kant believes that if we are to make any sense of the idea of a world of monads we must think of such a world as unified, for the idea of a world is not merely the idea of a multitude, but the idea of a whole. It is, then, worth looking briefly at Leibniz' problems with mind/body unity to understand what is at stake here.

It is well known that Leibniz believes that everything that really exists must be one, in the sense of being essentially unified. He is, however, also committed to the possibility of composite substances, with the composite being, "nothing but an accumulation or aggregate of simples" (*Monadology*, #2). The postulation of composite substances, however, presents Leibniz with a problem. For it is not clear how, given his own account of the nature of substance, something composite can really be an individual. For, it is not clear in what sense a composite of Leibnizian monads can be essentially unified. Many of Leibniz' contemporaries were committed to the coherency of the notion of composite

substance, because they believed that a human being is such a substance, being a substantial union, consisting of both mind and body. Of particular concern to Leibniz, and his contemporaries, then, was the attempt to give an account of the union of mind and body. It is not clear, however, whether Leibniz possesses the conceptual tools to explain the possibility of such a union. For Leibniz, or at least for the mature Leibniz, this question is no longer a question of the relationship between two distinct types of substance but the relationship between a single monad and a multitude of monads, for my mind is a monad, and my body also consists of monads. Leibniz' account of the unity of mind and body is to argue that my mind is 'dominant' over the monads that constitute my body and it is this relation of dominance that constitutes the (unifying) relation between my mind and (the monads that constitute) my body.¹⁷¹ Leibniz, however, denies the existence of real relations between monads, and so this dominance can only be explained in terms of the harmony between the individuals.

At least up until the late 1690s, Leibniz believed that the union between mind and body could be accounted for in terms of the harmony that existed between them. Thus, in his *New System of the Nature and Communication of Substances, and of the Union of Soul and Body* of 1695, Leibniz expounds his account of the nature of individual substances, and argues that,

There will be a perfect agreement among all of these substances, producing the same effect that would be noticed if they communicated through the transmission of species or qualities, as the common philosophers imagine they do. In addition, the organized mass, in which the point of view of the soul lies, being expressed more closely by the

¹⁷¹ This is why it makes sense to deal with Leibniz' account of the mind-body question in a Chapter on action. For the mind is dominant over the (monads that constitute) body by its **action** upon them.

soul, is in turn ready to act by itself, following the law of the corporeal machine, at the moment when the soul wills it to act; without disturbing the laws of the other – the spirits and blood then having exactly the motions that they need to respond to the passions and perceptions of the soul. **It is this mutual relation, regulated in advance in each substance of the universe, which produces what we call their *communication*, and which alone brings about *the union of soul and body*.**¹⁷²

Here, then, Leibniz argues that the unity of mind and body can be explained in terms of the harmony between them. However, as René Joseph de Tournemine pointed out in his *Conjectures on the Union of the Soul and the Body*, a work sympathetic to Leibniz and written in response to his *New System*, pre-established harmony does not seem strong enough to account for the *union* between mind and body. De Tournemine begins his article with a recounting of Leibniz's attack upon Descartes and the occasionalists. The Cartesians, de Tournemine argues, will tell you that,

the soul and the body. . . are united because to each change in the body there corresponds a change in the soul, and in the same way to each change in the soul there corresponds a change in the body.¹⁷³

And Tournemine praises Leibniz for pointing out that the mere correspondence between changes in the mind and changes in the body is not sufficient to account for real unity. He argues, however, that Leibniz's own position is subject to similar criticisms. Thus, he argues that Leibniz,

makes against the Cartesians an objection which entirely destroys their theory of the union of the soul and the body. Neither the law which God lays down for himself to act in parallel on the soul and on the body, nor the correspondence between the changes in the one and the changes in the other, can produce any genuine union between the soul and the body. There is, if you like, a perfect correspondence; but there is no real connection, any more than there would be between two clocks [the

¹⁷² Ariew & Garber (1989), pp.143-4 – my emphasis.

¹⁷³ Woolhouse (1997), p.247.

motions of which are perfectly matched]. There is no answer to this objection; but unfortunately, it destroys M. Leibniz's theory as well as that of the Cartesians. For after all, ***correspondence, or harmony, does not make a union, or essential connection.*** Whatever parallels we imagine between two clocks, even if the relation between them were perfectly exact, we could never say that these clocks were united just because the movements of the one corresponded to the movements of the other with perfect symmetry.¹⁷⁴

Leibniz himself accepts this criticism, and in a commentary on de Tournemine's article he recognizes that,

I have to admit that I would be greatly mistaken if I objected against the Cartesians that the agreement which, according to them, God maintains immediately between the soul and the body, does not create a genuine unity, because most certainly my *pre-established harmony* could not do any better.¹⁷⁵

Leibniz continues by suggesting that in offering his theory of pre-established harmony he "attempted only to give an explanation of the phenomena, that is to say, of the relation we perceive between the soul and the body" (*ibid.*). The metaphysical union between mind and body, however, is not phenomenal, and so Leibniz claims that he has "not taken it upon [himself] to look for an explanation of it" (*ibid.*). Leibniz's reaction here seems pretty disingenuous. For, it seems clear to me that Leibniz's monadology, of which his doctrine of pre-established harmony is an essential component, is clearly more than an attempt to 'explain the phenomena'. However, I am not primarily concerned here with providing an interpretation of Leibniz. Rather I am interested in drawing out a problem in Leibniz' metaphysics, a problem Kant attempts to solve. The problem Leibniz faces is that if the relationship between monads is merely that of harmony, it is difficult to see

¹⁷⁴ Woolhouse (1997), pp. 248-9 – my emphasis.

¹⁷⁵ Woolhouse (1997), p.250.

how a multitude of monads could possibly be unified. If we make the assumption that a 'world' must somehow be unified it is difficult to see how we can make any sense of a 'world' (or even 'a possible world') of monads. Some account of real interaction is required, then, if we are to make any sense of the idea of a world of individuals.

Although, as we have seen, Kant was not particularly concerned with explaining the relationship between mind and body because he believed that 'mind' and 'body' were not two ontologically distinct entities, he was concerned with the problems involved in conceptualizing composite individuals, for our idea of a world, or community, is the idea of something individual which is composed of individuals.

Chapter Eight

The Idea of Community and the Disjunctive Form of Judgment

I have argued that, for Kant, our idea of an intelligible world is the idea of a real community and that real interaction is only possible between individuals that constitute a real community. In this chapter I will examine in more detail what is involved in the idea of a real community and explain how it is related to the disjunctive form of judgment and the corresponding category of community. The main aim of this chapter will be to explain and justify the claim that, for Kant, the idea of a real community can be defined as the idea of a *real whole, the parts (or members) of which are simple and logically prior to the whole and which mutually limit or resist one another*. In the following three sections of this chapter I will examine Kant's analysis of the idea of community in more depth.

In (8a), I will examine his account of **the concept of community**, introduced in the table of categories in the *Critique of Pure Reason*. I will explain how this category is related to the disjunctive form of judgment and will argue that **the category of community is the concept of a whole the parts (or members) of which mutually exclude one another**.

In addition I will explain what Kant means in claiming that the concept of resistance is a predicable of the category of community. This claim, that Kant makes in passing, is often overlooked by commentators. However, given the role the concept of resistance plays in his model of action, this claim is highly significant, for it explains why and how Kant believes that action (and interaction) is intelligible.

In (8b), I will explain how **the idea of community** is to be distinguished from the *concept*. I will begin by examining his general account of the distinction between concepts and ideas and I will argue that the idea of community is distinguishable from the category of community in, at least, three ways: **in the case of the idea of a community: (i) its parts/members must be logically prior to the whole, (ii) its parts/members must be simple (i.e. they cannot themselves have parts) and, (iii) it must be an absolute whole, in the sense of being a whole that is not itself part of any other whole.** I will also explain why, given this definition of ‘community’, phenomenal objects cannot be experienced as members of such a community. In other words I will explain why nothing given in experience can be experienced as an object corresponding to our idea of a member of (such) a community.

In (8c), I will explain the distinction between the idea of an **ideal** community and that of a **real** community. Our idea of a **real** community is **the idea of a real as opposed to an ideal whole.** This distinction has to do with the nature of the *unity* of the whole. The distinction can be expressed in mentalistic terms. An ideal whole is a whole whose unity exists merely in the mind of the observer, say God. A real whole, on the other hand, is a

whole whose unity is *intrinsic* to the whole, in the sense that the whole itself is the source of its unity. For Kant, a community is unified by inter-substantial laws, where laws are thought of as intelligible relations that bind the members of the community together. In the case of an ideal community there will be a harmony between the change of state of one substance and that of another. An ideal observer could recognize regularities between the change of state of one substance and that of another and could formulate inter-substantial laws to capture these regularities. The individual substances themselves, however, are not governed by these laws, nor are they the source of the laws, for the laws only exist in the mind of the observer, and so although the observer experiences the individuals as a whole they are not, in themselves, a whole as there is nothing that unifies them. In a real community, on the other hand, the community itself must be the source of the laws. Now, as the community just is its members, then the members of the community themselves must be the source of the laws that provide the community with its unity. This is why a real community must consist of autonomous agents. For an autonomous agent is, by definition, an individual that gives (i.e. is the source of) laws.

(8a) The Category of Community

In this section I will argue that, for Kant, the concept of community is the concept of (a) a **whole**, (b) the parts/members of which **mutually exclude** one another. I will explain how Kant believed that this concept is derived from the disjunctive form of judgment.

In the *Critique of Pure Reason*, Kant introduces the category of community as the third category of relation. The structure of the table of categories is derived from the table of judgments, and this table is divided into four classes, into judgments of quantity, of quality, of relation and of modality. The categories of the third class, then, are derived from the judgments of relation. According to Kant there are three types of relational judgment: categorical judgments (A is B), hypothetical judgments (if p then q) and disjunctive judgments (p or q or r). The categories of substance and accident are derived from the categorical form of judgment. The categories of cause and effect are derived from the hypothetical form of judgment and the category of community is derived from the disjunctive form of judgment. The table of categories leads to a table of principles, with the principles being, “nothing other than rules of the objective use of the categories” (A161/B200). Thus, the principles can be thought of as rules for the application of the categories to (empirical) objects. The analogies of experience are the third set of principles and correspond to the categories of relation – with each of the analogies being presented as an account of why one of the categories must necessarily be applied to objects of experience. Thus the first analogy concerns the application of the category of substance, the second that of causality and the third that of community.

Kant believes that the category of community (and as a result the notion of interaction) is to be sharply distinguished from that of cause and effect, for they are derived from different forms of judgment. We understand the importance of this claim by considering an alternative way of conceptualizing interaction. Defenders of such an alternative conception of interaction would argue that we can fully capture what is involved in

interaction in the following terms: when two entities, say x and y , interact x has a causal relation to y and y has a causal relation to x . Kant does not deny that this partially captures what is involved in the relation of interaction¹⁷⁶, but he does not believe that it is the full story¹⁷⁷, for he believes that when a number of entities interact they (a) constitute a **whole** and (b) mutually **exclude** one another. These two factors are essential to the relation of interaction and cannot be captured by appealing to the ideas of ground and consequence or to the hypothetical form of judgment. Thus, in his commentary to the table of categories in the *Critique of Pure Reason* Kant compares the causal relation to the relation of interaction/community and points out that in the case of simple causation the relation is one of **subordination**, whereas in the case of interaction the relation is one of **coordination** (B112). What he means by this is that in a causal relation the consequence is subordinated to the ground. For this reason the ground-consequence relation is the principle of the series, for the relation of ground and consequence can provide us with a well ordered chain of causes and effects. The relation of community, on the other hand, cannot be understood in terms of the idea of subordination, for when a number of entities are members of a community they are not subordinated to one another but are coordinated with one another. The concept of co-ordination cannot be understood in terms of mutual subordination. When entities are coordinated with one another they are parts of a whole and mutually exclude one another.

¹⁷⁶ “[T]he third category always arises from the combination of the first two in its class” (B110). In the case of the category of relation, which is the third category of relation, the first and second categories are substance and causation. So community involves substances in causal relations, but cannot be reduced to the notion of mutual causation.

¹⁷⁷ “But one should not think that the third category is therefore a merely derivative one and not an ancestral concept of pure understanding. For the combination of the first and second in order to bring forth the third concept requires a special act of the understanding, which is not identical with that act performed in the first and second” (B111).

Thus Kant explains that the relation of community/interaction,

Is an entirely different kind of connection from that which is to be found in the mere relation of cause to effect (of ground to consequence), in which the consequence does not reciprocally determine the ground and therefore does not constitute a whole with the latter (as the world-creator with the world). The understanding follows the same procedure when it represents the divided sphere of a concept as when it thinks of a thing as divisible, and just as in the first case the members of the division exclude each other and yet are connected in one sphere, so in the later case the parts are represented as ones to which existence (as substances) pertains to each exclusively of the others, and which are yet connected in one whole. (B113)

In the first sentence of this passage Kant distinguishes the concept of causation from that of interaction, and focuses on the fact that in the case of interaction the entities in interaction “constitute a whole”. This is not the case in the ground-consequence relation. He appeals to the example of God, the “world-creator”. God is the ground or cause of the world, but God and the world do not constitute a whole. If God were thought of as interacting with the world, however, God and the world would constitute a whole.

To understand the second sentence of this passage it is necessary to have a closer look at Kant’s account of the disjunctive form of judgment. A disjunctive judgment has the form: ‘x is A or B or C’.¹⁷⁸ Kant explains this form of judgment in the *Critique of Pure Reason* in the following terms: “in all disjunctive judgments the sphere (the multitude of everything that is contained under it) is represented as a whole divided into parts (the subordinate concepts)” (B112). He makes his point a little more clearly in his logic lectures. In his *Jäsche Logic*, for example, he give the following explanation:

¹⁷⁸ Or perhaps more accurately: “x is A or x is B or x is C”.

disjunctive judgments represent various judgments as in the community of a sphere and produce each judgment only through the restriction of the others in regard to the whole sphere. (9:107)

A disjunctive judgment, then, is judgment in which a number of judgments somehow restrict one another and fill up a (logical) sphere. To understand what Kant means by this it is instructive to look at an example. In his *Jäsche Logic* Kant himself gives the following example of a disjunctive judgment: “A learned man is learned either historically or in matters of reason.” (9:108) Here the concept ‘learnedness’ is divided into ‘parts’. The concept ‘learnedness’ is in this case the logical ‘sphere’ that is to be divided into parts. The parts of this sphere are ‘learned historically’ and ‘learned in matters of reason’. These parts mutually exclude one another in the sense that in so far as one is ‘learned historically’ one is not ‘learned in matters of reason’ and, Kant believes, taken together they completely ‘fill the sphere’ of the concept of learnedness in the sense that they exhaust the concept. In other words, Kant maintains that the ‘or’ in a disjunctive judgment is an exclusive ‘or’, and that in such a judgment the members of the disjunction exhaust the concept. In the disjunctive judgment, then, we find a number of judgments mutually excluding one another and completely filling a logical space. This conception of a logical ‘space’ allows us to think of a ‘space’ that has parts but which is not, unlike the space of intuition, infinitely divisible. This allows us to think of a whole the parts of which are simple. This will be important when we turn to the idea of community.

It is, then, from the disjunctive form of judgment that we get the concept of ‘exclusion’. Kant makes this clear in his commentary to the table of categories. In this section he

compares the disjunctive form of judgment with the hypothetical (if. . . then) form of judgment, and asks us to,

note that in all disjunctive judgments the sphere (the multitude of everything that is contained under it) is represented as a whole divided into parts (the subordinate concepts), and since none of these can be contained under any other, they are thought of as **coordinated** with one another, not subordinated, so that they do not determine each other **unilaterally**, as in a **series**, but **reciprocally**, as in an **aggregate** (if one member of the division is posited, all the rest are excluded, and *vice versa*. (B112)

Earlier in his commentary on the table of categories, Kant explains that the categories he has listed do not provide a complete list of the *a priori* concepts of the understanding, for there are also derivative concepts, which Kant calls “predicables”, which can be derived from the categories. Thus, Kant explains that

For the sake of the primary concepts it is therefore still necessary to remark that the categories, as the true **ancestral concepts** of pure understanding, also have their equally pure **derivative concepts**, which could by no means be passed over in a complete system of transcendental philosophy, but with the mere mention of which I can be satisfied in a merely critical essay. (A81-2/B107)

Under the category of community Kant lists two “derivative concepts” or predicaments: presence and resistance (A82/B108). The reason why resistance is a predicable of the category of community is because our (pure, unschematized) concept of resistance is to be understood in terms of exclusion, and we understand the notion of exclusion *a priori* through our grasp of the disjunctive form of judgment. What we mean if we claim that one thing resists another is that if (or, insofar as) the thing is posited all the rest are excluded.

In the previous chapter I explained Kant's theory of action. On this account, one individual acts upon another if the agent withdraws some resistance which allows a dead power (conatus) in the patient to become a living power. The possibility of action, then, presupposes resistance, and I claimed that Kant believes that individuals can only resist one another if they are members of a community. We are now in a better position to understand this claim for Kant believes that resistance is a **pure** concept parasitic upon the category of community. One individual resists another individual by excluding it from a "space". The metaphor I have just used is spatial, and if our conceptualization of resistance/exclusion necessarily relied upon such spatial metaphors then the concept of resistance would not be a pure concept. Kant's whole point, however, in arguing that resistance is a predicable of the category of community, is that the notion of mutual exclusion, and the related notions of resistance (and impenetrability) although they can be applied to phenomenal objects in space, are **pure** concepts derived from the disjunctive form of judgment, and as such can be thought independently of their conditions of application to objects given in intuition. In other words, Kant believes that we *understand* the notion of one individual excluding another from a 'space' without any appeal to intuitive space, for our understanding of the notion of the 'space' or 'sphere' of a concept is logical and not intuitive. The category of community, then, allows us understand the notion of a number of impenetrable individuals (concepts) filling a conceptual space (another concept) and excluding other individuals (concepts) from their bit of the conceptual space, without any appeal to the space of intuition. This is the basis for our capacity to think of disembodied spirits (i.e. non-spatio-temporal intelligible individuals) as interacting.

(8b) - The Distinction Between the *Category of Community* and the *Idea of a Community* (or ‘World’)

In this section I will examine Kant’s distinction between the category of community and the *idea*. Kant argues that whereas our categories are always conditioned our ideas are concepts which are unconditioned. The *concept* of a community is the *concept* of a whole, the parts/members of which mutually exclude one another. The idea of a community is unconditioned in the sense that its parts/members must be (a) **simple** and (b) **logically prior to the whole**. And the whole must be (c) **absolute**, in that it must be a whole that is not part of any other whole. These three facts about the idea of community are the reason why nothing in experience can be experienced as a member of a community. For, given the spatio-temporal nature of our experience, nothing that is experienced can be experienced as simple, and in every experience the whole (of space) is logically prior to part. As I am primarily concerned with Kant’s understanding of the idea of a member of a community, I am not, at least in the context of this dissertation, particularly interested in Kant’s claim that a community must be an absolute whole, and so, in what follows I shall focus exclusively on his claim that the members of a community (corresponding to our idea as opposed to our mere concept of a community) must be simple (i.e. individuals) and logically prior to the whole.¹⁷⁹

¹⁷⁹ Kant makes it clear that our idea of a world (which is the idea of a community) must be the idea of an absolute whole in his metaphysics lectures. Thus he explains that, “A composite <*compositum*> can be either a relative <respective> or an absolute whole <*absolute totum*>. It is a relative whole <*totum respective*> insofar as it is not a part of a whole of the same kind, but **an absolute whole** insofar as it **is a part neither of the same nor of another kind**, e.g., a house is a relative whole insofar as it is a whole of its kind; but it is not an absolute whole, for it is a part of another kind, namely of a street. // **The world is**

As our pure idea of a ‘world’ (the ‘intelligible world’) is the idea of a community in this sense, in the following sections I will treat the expressions ‘idea of a community’ and ‘idea of a world’ as synonyms.

Categories and ideas

I have argued that, for Kant, the intelligible world is the pure *idea* of a community, the members of which must be simple and logically prior to the whole. To understand why the idea of a member of a community must have these features we must understand how Kant distinguishes between concepts and ideas in general. Kant’s most famous account of the distinction is to be found in the *Critique of Pure Reason*. Here Kant offers two accounts of the distinction between concepts and ideas. **First**, he distinguishes between the understanding, as the source of *concepts* and reason as the source of ideas. The understanding is the faculty of judgments, whereas reason is the faculty of syllogisms. **Second**, he argues that whereas concepts are conditioned, ideas are always unconditioned. In addition, in his lectures he offers a **third** account of this distinction, arguing that ideas “come about when one enlarges a concept of the understanding to infinity” (29:848). I believe that the second and third accounts of the distinction are the most useful.

no relative whole, but rather **an absolute whole** in the metaphysical sense” (*Metaphysik Mrongovius*, 29:851 – my emphasis).

In his lectures on metaphysics Kant explains the distinction between concepts of the understanding (concepts/categories) and concepts of reason (ideas) in the following terms:

We have now two sorts of concepts in our soul: concepts of understanding and of reason. Concepts of reason [i.e. ideas] come about when one enlarges a concept of the understanding to infinity. (*Metaphysik Mrongovius*. 29:848)

To understand what Kant might mean by ‘enlarging a concept to infinity’, we must understand what Kant means by infinity. Kant consistently distinguishes between ‘real infinity’ and ‘mathematical infinity’. Mathematical infinity “arises through the successive addition of one to one” (28:569). Given Kant’s understanding of the nature of mathematics a pure concept of the understanding is not the sort of thing that could be ‘enlarged to infinity’ in the mathematical sense. For numbers, Kant believes, are constructed in intuition, and so are not pure. Mathematical notions, such as the notion of mathematical infinity, then, cannot be applied to pure concepts.¹⁸⁰ This is not the only way to understand ‘infinity’, for Kant explains that,

Infinite can actually be taken in two senses. In the first the concept of the infinite is a pure concept of the understanding, and then it is called: real infinity <*infinitum reale*>, i.e., in which there are no negations. (28:569. *Metaphysik L2*)

When Kant talks of ‘enlarging a concept to infinity’ he must mean ‘infinity’ in this sense. So we can say that an idea, being a category “enlarged to infinity”, is “a concept in which there are no negations”.

¹⁸⁰ In addition, Kant believes that there is no mathematical infinite. He agrees that it is true that “beyond every number I can add a higher one”, but denies that there is such a thing as “the largest number” (*ibid.*). To claim that space, for example, is (mathematically) infinite is merely to claim that, “the concept of the magnitude of space is never total” (*ibid.*).

An idea must be 'infinite', in the sense that it must be based upon what I shall call a *perfect* judgment. A *perfect* judgment is a judgment that contains nothing negative. Examining Kant's account of the real, as opposed to the mathematical, infinite we can understand why a concept enlarged to infinity, that is one representing a perfect judgment, must contain nothing negative. For in a perfect disjunctive judgment each member of the disjunction must itself be a positive judgment. Now the category that corresponds to the positive judgment is the category of reality. Thus, as a result, each member of a community (corresponding to a perfect disjunctive judgment) must be *real*.¹⁸¹

Both the category and the idea of a world are ultimately derived from the disjunctive form of judgment. A disjunctive judgment can, in principle, contain negative judgments. For example, 'A is x or A is not-x'. A disjunctive judgment 'enlarged to infinity' would be a disjunctive judgment which contains no negations. A perfect disjunctive judgment, then, would have the form 'A is x or A is y or A is z'. In such a judgment the quality of each member of the disjunction is affirmative. Now the category which is derived from the affirmative form of judgment is that of reality. Thus, the claim that the idea of a world must be derived from a perfect disjunctive judgment implies that in the idea of a world each member (or part) of the community must be 'real'.

The simplicity of members of a world

¹⁸¹ And this is one reason why in our idea of a world, the members must be logically prior to the whole, for in a perfect judgment each element of the judgment must be real and logically precede the judgment.

The fact that Kant believes that the idea of a member of a world must be simple is not difficult to demonstrate. I suggest that Kant offers at least two distinct arguments for the claim that in our pure idea of a community its members (or parts) must be thought of as simple. The first argument is based upon the assumption that the idea of a community is the idea of a real composite. The second is based upon the claim that “reason demands completeness” and anything divisible is not complete. The fact that our idea of member of community must be the idea of something simple is one reason why Kant believes that no object of experience can be experienced as a member of (such) a community.

In his lectures Kant distinguishes between the notions of an ideal and a real composite, explaining that,

The concept of a composite presupposes parts. When the parts of a composite can be given prior to the composition, then it is a real composite. But when they cannot be given prior to the composition, then it is an ideal composite. (*Metaphysik L2*, 28:565)

Elsewhere, he argues that,

Simple substance is called a monad. . . Considered as noumenon, the world certainly consists of simples, for composition is just mere relation. But in the world of appearances, there are no simple parts. Only the intelligible world, noumenal world, is of monads, but we do not at all cognize it. (*Metaphysik Dohna*. 1792-3. 28:663-4)

The argument he provides here is based on the fact that our idea of a world is the idea of a composite, and composition requires (simple) things that are composed. He repeated this argument frequently. For example, in his lectures from the early 1780s he argues

that, “The world, considered as noumenon, must consist of simple parts, because otherwise it cannot be composed” (Metaphysik Mrongovius, 29:859). Similar passages are not hard to find.¹⁸² This argument is very similar to Leibniz’s (1989) argument at the start of the *Monadology* that, “there must be simple substances, since there are composites; for the composite is nothing more than a collection, or aggregate, of simples” (p.215). Unfortunately, it is not clear how this follows from the fact that our idea of the intelligible world is the idea of community, for the category of community does not involve the notion of composition¹⁸³, and so if this argument is to be successful Kant must explain why our idea of community (as opposed to the concept) must be the idea of a composite. Although Kant himself doesn’t explicitly offer an argument it is clear that he believes that our *idea* of a community is necessarily the idea of a real composite. This is not the case with the *category* of community, which can equally well be applied to non-composite wholes, such as space. For space, Kant argues in the third analogy, is to be thought of as a community, space however is not a real composite, for it is not composed of spaces; rather, (particular) spaces are introduced by introducing limitations into space.

One reason for this is provided by the fact that he believes that an idea “is a concept enlarged to infinity”, and we have seen that what this means is that an idea can contain nothing negative. As a result of this our *idea* of a part cannot be introduced by

¹⁸² For example: “Now the question arises: can one say of every substantial composite <*compositum substantiale*> that it consists of simple substances, i.e., it is a whole consisting of monads <*monadatum*>? Yes, insofar as it is *noumenon*, for all connection is nothing other than relation. Since the substances by definition <*ex definitione*> are privy to outer existence for themselves, one can remove all relation and the substances remain and are simple” (Metaphysik Mrongovius. 1782-3. 29: 827).

¹⁸³ We think of space as subject to the category of community, but space is not a composite of spaces.

introducing limitations into a whole, for to introduce limitations is to introduce something negative. If in our idea of a community the parts cannot be introduced by introducing limitations into a whole, then the parts must be logically prior to the whole, and so in our idea (as opposes to mere concept) of a whole the whole must be thought of as composed of parts, rather than the parts being thought of as being introduced into a pre-existing whole through the introduction of limits.

Although Kant believes that the fact that a world must be thought of as a real composite implies that its parts must be thought of as simple, his argument is not compelling, for the most he can show is that in a real composite the parts must be logically prior to the whole. However, just because the parts of a real composite cannot be introduced by introducing limitations or divisions into the whole, we cannot draw the conclusion that the parts themselves can't be thought of as divided.

Elsewhere Kant offers what seems to be an independent argument for the claim that in our idea of a community the parts of the community must be simple (in the sense of being indivisible). The crux of this argument is that ideas are the product of reason, and reason demands completeness. Thus, for example, he argues in his logic lectures that,

The concept of a part is a *conceptus purus intellectualis, seu notio*. But the concept of a part that is not composite is a *notio rationis*, idea. As long as my reason represents something divisible, that can always be divided further. But my reason finally demands the ultimate part, which cannot be further divided into parts, i.e., is simple. This concept cannot be shown in experience, and this is a concept *a priori*, or idea. (*Vienna Logic*. 907)

Here he argues that the concept of a part is a pure concept, and the concept of a simple part is an idea (*notio rationis*). Reason demands completeness, and the idea of a community whose members (or parts) were not simple would not be complete. It is also arbitrary, for there is no reason for reason to stop the division at this particular point. It seems to me that this argument is the real basis for Kant's claim that the members of an intelligible world must be thought of as simple (indivisible) individuals.

Phenomenal community and the third analogy

This explains why Kant believes that no object of experience can be experienced as a member of a world. The reason for this is that it is an essential feature of phenomenal space that the whole of space is logically prior to the parts of space. For a part of space is always thought of as being 'in' a greater space. The reason for this is because the parts of space are only brought into existence by introducing limitations into this greater space. Thus the concept of a 'part' of space logically presupposes a greater space. Thus, Kant argues in the *Transcendental Aesthetic* that it is a defining characteristic of space that the whole precedes the parts. As Kant explains:

One can only represent a single space, and if one speaks of many spaces, one understands by that only parts of one and the same unique space. And these parts cannot as it were precede the single all-encompassing space as its components (from which its composition would be possible), but rather are only thought in it. It is essentially single; the manifold in it, thus also the general concept of spaces in general, rests merely on limitations (A25)¹⁸⁴

¹⁸⁴ Here Kant makes it clear that in the case of space the whole logically precedes the parts, for the parts rest upon limitations of the whole. The recognition of this feature of space is what motivates Kant's claim that our representation of space cannot be conceptual and must as a consequence rest upon pure intuition.

The idea of a community is based upon the idea of a perfect disjunctive judgment. A perfect disjunctive judgment (which is the basis of our *idea* of community) is infinite in the sense that each term in the disjunction is (transcendentally) affirmative. In an imperfect disjunctive judgment (the basis of our category of community) some of the terms are (transcendentally) negative. The judgment “x is either mortal or not-mortal” is imperfect in this sense. This helps us understand why although everything in space is subject to the category of community, nothing appearing in space can be experienced as being a member of an intelligible world (community). We never experience space itself, but always only a part of space. All spatial experience, then, involves an awareness that there is some space that is not being experienced, that the space we are experiencing is a particular space. Whenever we experience anything in space we are aware that there is some (surrounding) space that we are not experiencing that limits the space we are experiencing. We do this, Kant believes, by thinking of the space we are experiencing as a member of a community of spaces, for the relation between spaces is a logical relation. However, this judgment is always imperfect, for it is an irreducible aspect of spatial experience that we judge that: “the whole of space is either present to me or not present to me”. We are aware that the part of space we are intuiting is part of space as a whole through such an imperfect disjunctive judgment. This is what lies behind Kant's claim that “the general concept of spaces in general rests merely on limitations”, for a limitation is something negative which has to be introduced through an imperfect disjunctive judgment.

(8c) The Distinction Between the Ideas of a Real and an Ideal Community

The idea of a community is the idea of a whole, the parts (or members) of which are simple and logically prior to the whole and which mutually limit or resist one another. We can, however, distinguish between our ideas of a real and an ideal community. This difference can be explained in terms of the distinction Kant draws between a real and an ideal whole. The difference between these two types of whole has to do with the nature of the unity of the whole. A real whole is, in scholastic terminology, a *unum per se*, whereas an ideal whole is merely a *unum per accidens*. Kant believes that a composite individual can only be a real whole if the individuals that constitute the whole are the source of the unity of the whole. Our idea of a real community, then, is the idea of a *real* whole, the parts (or members) of which are simple and logically prior to the whole and which mutually limit or resist one another.

Thus, Kant explains in his lectures on metaphysics of 1790-1 that,

Substances are the matter of the world, the formal aspect of the world consists in their connection <*nexu*> and indeed in a real connection <*nexu reali*>. The world is thus a real whole <*totum reale*>, not ideal. (*Metaphysik L2. 28:581*)

Our idea of a world is the idea of a *real* as opposed to an *ideal* whole. Elsewhere in the same lectures, Kant is a bit more explicit about this distinction. He argues that,

The connection <*nexus*> is ideal if I merely think the substances together, and real if the substances actually stand in interaction <*commercio*>. // The form of the world is a real connection <*nexus realis*> because it is a

real whole <*totum reale*>. For if we have a multitude of substances, then these must also stand together in a connection, otherwise they would be isolated. Isolated substances, however, never constitute a whole <*totum*>, then they must also be a real whole <*totum reale*>. For were they ideal, then surely they could be represented in thought as a whole <*totum*>, or the representations of them would constitute a whole <*totum*>; but things in themselves would still not constitute a whole on this account. (*Metaphysik Mrongovius*, 29:851)

An ideal whole is a whole that can be “represented in thought” as a whole. In such a whole the unity only exists in the mind of the observer. In a real whole, in contrast, the unity must be intrinsic to the whole. Although Kant himself does not explicitly make this claim, I suggest that what this means is that the individuals that constitute the whole must be responsible for the unity of the whole. In the case of a community, what unifies the whole are laws, and so the individuals must be thought of as the source of these laws which provide the community with its unity. In the following chapter I shall examine the nature of these laws in more detail.

Conclusion

To summarize: The idea of a community is the idea of a whole, the parts (or members) of which are simple and logically prior to the whole and which mutually limit or resist one another. We can, however, distinguish between our ideas of a real and an ideal community. This difference can be explained in terms of the distinction Kant draws between a real and an ideal whole. Our idea of a real community, then, is the idea of a *real* whole, the parts (or members) of which are simple and logically prior to the whole and which mutually limit or resist one another.

A real community is a real whole whose parts are simple but connected. As a real whole it is an individual (in the sense of being a *unum per se*) whose parts/members are individuals. As a result something might fail to be a real community if: (a) It is an individual but its parts are not themselves really individuals. Or: (b) Its parts/members are really individuals but it is not itself really an individual. The phenomenal 'world' fails to be a real community (or world) for the first reason. The idea of a 'world' of isolated Leibnizian monads fails to be the idea of a real community (or world) for the second reason.

Chapter Nine

The Civil Condition as a Real Community:

An Analysis of Kant's Ontology of Property in the *Doctrine of Right of the Metaphysics of Morals*

In this concluding chapter I will attempt to bring together a number of claims made in the two halves of this dissertation. In the first part of the dissertation I argued that the idea of a good will or autonomous individual is the idea of a member of an ideal community, and that to be virtuous is to strive to be such an individual. In the second part I argued that for Kant the pure idea of a community is the idea of a real whole the members (or parts) of which are simple and logically prior to the whole and which mutually limit or resist one another. I also argued that such resistance is necessary for interaction between individuals. In this chapter I will argue that Kant believes that our idea of a political community governed by juridical laws (or what Kant often calls the “civil condition”) is the idea of a community in this sense. Indeed, I suggest that Kant believes that the **only** way we can conceive of a real community is as a political community. If I am right then Kant believes that the idea of a kingdom of ends is the idea of an ideal political community, and that to be virtuous is to strive to be a member of such an ideal political community.

The bulk of this chapter is devoted to examining Kant's arguments in the *Doctrine of Right* of the *Metaphysics of Morals*. This is a long and complicated text, and so at most all I can offer in such a short space is the outline of a strategy for reading it. I should note at this point that I believe that in the *Doctrine of Right* Kant is attempting to do two very different things. On the one hand he is attempting to describe the principles of an ideal political community. On the other hand he believes that we should not only strive to be potential members of such a ideal (holy) community but that we should strive, in so far as it is possible, to bring the societies we live in closer to this ideal. As a result of this, his second goal in the *Doctrine of Right* is to examine to what degree, and how, human communities can be brought closer to this ideal, for Kant believes that it is impossible for human societies to instantiate this ideal (at least without divine intervention). In other words, I suggest that Kant often shifts between attempting to describe the principles of an ideal political community as such, and the principles of an 'ideal' human community. Kant is working in the natural law tradition¹⁸⁵, and one could say that he believes that an ideal community of rational beings would be governed by natural law, whereas an 'ideal' human community can only be governed by positive law, and that because of the imperfections of our natures positive law is necessarily distinct from natural law. Although a thorough discussion of the relationship between natural law and positive law is beyond the scope of this dissertation, it is worth pointing out at this juncture. For it lays behind my general strategy for reading the *Doctrine of Right*. In what follows I attempt to explain Kant's account of an ideal political community, that is a community governed by principles of natural right and law and my general strategy for explaining those passages

¹⁸⁵ For a convincing illustration of this see Gregor (1993).

that seem to contradict my interpretation would be to argue that in these passages Kant is talking about an (ideal) human society, a society governed by positive statutory, as opposed to natural, law.

(9a) Introduction

I have argued that Kant argues: (a) An individual *a* acts upon another individual *b* if *a* withdraws some impediments which allows a change to occur in *b*. According to this model of action, the patient is the ground of both determinations, but the agent is the ground of the change of determinations. (b) Real interaction is only possible between members of a real community. And, (c) a real community is only possible if each individual member of the community has ‘given’ the laws that unify the community.

This account of interaction is highly abstract. In this chapter I will offer a concrete illustration. Kant's account of property in the *Doctrine of Right* of the *Metaphysical of Morals* is based upon an analysis of the ideal of “the civil condition” and I will show that Kant conceives of the ideal of a “civil condition” as community understood in these terms, and that his account of property (and in particular his account of the transferal of property) has to be read in the context of his metaphysical analysis of interaction and the idea of community.

Showing that Kant believes that our idea of the civil condition is based upon the idea of a community is the weak thesis of this chapter. Even if I only manage to convince the

reader of this, I believe I would have demonstrated something interesting. Ultimately, however, in this chapter I wish to argue for a much stronger thesis. Namely, the claim that Kant believes that our pure idea of community **is** the idea of an ideal civil condition. In other words, I will argue that Kant believes that the **only** way we can conceive of a community is as a civil community. In arguing this I am rejecting Korsgaard's influential interpretation of Kant's project. For, in her *Creating the Kingdom of Ends* she argues that Kant conceives of the kingdom of ends as a non-political community, analogous to an ideal community of friends. I argue, in contrast, that the idea of a kingdom of ends is the pure idea of a community and that Kant believes that the only pure idea of a community we possess is the idea of a civil condition, and that this is the idea of a political community governed by juridical laws. Juridical laws are laws that assign rights and corresponding duties. Such laws are coercive, Kant believes, in that the possession of a right always implies duties in others. If I have a legitimate right to something I can legitimately demand that others do not interfere with my possession of it. In other words, juridical laws are coercive, in that that they permit individuals to legitimately **resist** one another, for juridical laws assign legitimate rights, and to have a legitimate right is to be empowered by a law to resist another individual in their use of an object. Now, given Kant's belief that the only way one individual can act upon another is by withdrawing resistance, the only type of laws that can be the basis of interaction are juridical laws. Only such laws, then, can be the basis of a community in which there could be any interaction. In addition juridical laws are the only type of laws that create reciprocal relations between individuals, and a community can only be unified if the relations

between the members of the community are reciprocal. As a result, we cannot even conceive of a community of individuals in interaction that is not a civil community.

In contrast to Korsgaard, and to Swedenborg, who conceived of heaven as a community of spirits governed by laws of love, Kant, at least by the 1780s, maintains that it is impossible for us to even conceive of a community governed solely by laws of benevolence or laws of love. This is not to say that Kant denies that we have a duty of benevolence, it is just that he believes that benevolence itself is only possible in the context of a civil condition.¹⁸⁶ The reason he believes this is because he does not believe in any natural property, instead he believes that property is only intelligible in the context of a system of juridical laws and corresponding legitimate rights, and as a consequence of this he believes that without such a system of laws no one would be able to possess anything that they were capable of giving. Thus he claims in his lectures that that, “If all men were willing to act from benevolence merely, there would be no ‘mine’ and ‘thine’ at all, and the world would be a stage, not of reason, but of inclination” (*Ethik Collins*, 27:416). The idea of a community governed purely by laws of benevolence with no laws determining rightful ownership, then, is not an idea of reason. The idea of such a community is self-contradictory, because no one would have anything they could legitimately give¹⁸⁷ and no one would be able to act intelligibly upon anyone else because there would be no intelligible resistance to withdraw.

¹⁸⁶ This claim will be modified later in this chapter. But it will do for now.

¹⁸⁷ And no one would be able to receive.

If my interpretation is correct it implies a radical re-evaluation of the relationship between the *Doctrine of Right* and the *Doctrine of Virtue*. On the standard reading of Kant's moral philosophy his ethics is more basic than his 'political philosophy'. Thus Rosen (1993) remarks that, "for a long time it was fashionable to regard Kant's political writings as minor works" (p.1). If I am correct, however, Kant's political philosophy stands at the heart of his ethics because the idea of a kingdom of ends is, and can only be, conceived of as a political community governed by juridical laws.¹⁸⁸ For this reason Kant's ethics is based upon his so called 'political philosophy', for his doctrine of rights provides his ethics with its content. This is why the *Doctrine of Right* precedes the *Doctrine of Virtue*, for to be virtuous is to strive to be a member of an ideal political community governed by laws of right.

In chapter one I explained Kant's distinction between virtue and holiness and argued that although (the possibility of) virtue is the *ratio cognoscendi* of holiness, holiness is the *ratio essendi* of virtue, for what it is to be virtuous is to strive for holiness. In this chapter I will argue that Kant believes that the principle of right should be regarded as a principle of holiness. What I mean by this is that the principle of right, unlike the principle of virtue, is a principle that can govern rational beings as such (including holy beings), and not merely beings such as us who have a sensuous nature.¹⁸⁹ As a result of this the

¹⁸⁸ Kant's political philosophy provides his ethics with its content, or, at least, with *most* of its content.

Later in this chapter I shall discuss the duty of benevolence, which, Kant believes is a purely ethical duty.

¹⁸⁹ Thus Kant explains in the Introduction to the *Doctrine of Virtue*, that the concept of (ethical) duty is the concept of "necessitation" or "constraint" and that "such constraint. . . does not apply to **rational beings as such** (there could also be **holy** ones) but rather to *human beings*, rational *natural* beings, who are **unholy enough** that pleasure can induce them to break the moral law even though they recognize its authority; and even when they do obey the law, they do it *reluctantly* (in the face of opposition from their inclinations), and it is in this that such *constraint* properly consists" (6:379-80). But he adds in a footnote that, "*as a*

principle of right should be regarded as the *ratio essendi* of the principle (or, as we shall see, principles) of virtue.¹⁹⁰

In the *Groundwork* Kant begins by examining our everyday understanding of what it is to be virtuous in order to discover what the virtuous person is striving to be, and he reaches the conclusion that the virtuous person is striving to be a member of an ideal kingdom of ends. In the *Metaphysics of Morals* Kant goes in the other direction, starting by examining the principle of right before examining the principle of virtue. The principle of virtue is a principle that is not applicable to holy beings, for to be virtuous is to strive towards holiness and so, by definition, a holy being could not be virtuous. The Doctrine of Virtue (and the principles and duties of virtue), then, are only applicable to sensuous agents, such as human beings. In the *Doctrine of Right*, however, Kant is not (primarily) attempting to describe an ideal human community, but an ideal civil condition; the principle of right, then, is supposed to present us with a principle of morality that is valid for all rational beings. The doctrine of virtue, on the other hand, examines how this moral ideal presents itself to imperfect beings such as we are: beings with needs, who can only experience ourselves as existing in space and time.

Thus, in the *Metaphysics of Morals*, Kant explains that, “for finite holy beings (who could never be tempted to violate duty) there would be no doctrine of virtue but only a

moral being [a human being] is also **holy enough** to break the moral law *reluctantly*” (ibid. – my bolding throughout).

¹⁹⁰ The principle of right can be called a principle of holiness, for it is a principle that could govern an ideal community of holy beings. The principle of virtue, in contrast, is a principle only applicable to imperfect beings such as we are, beings who can, at the most, only strive towards holiness.

doctrine of morals” (6:383). As Kant normally contrasts the doctrine of virtue with the doctrine of right, this passage suggests that he regards the doctrine of right as a doctrine of morals that is applicable to rational beings as such, including “finite holy beings”. I argue, then, that the doctrine of right should be regarded as providing an account of the nature of the laws that could govern a (political) community of finite holy beings. We have a duty to become a member of such a (political) community, and the doctrine of virtue is an account of what is involved in striving to be a member of such a community for creatures such as we are. Kant makes it clear that he conceives of the kingdom of ends as an ideal state in the *Religion*, where he claims that,

The teacher of the Gospel manifested the Kingdom of God on earth to his disciples only from its glorious, edifying and moral side, namely in terms of **the merit of being citizens of a divine state**; and he instructed them as to what they had to do, not only that they attain to it themselves, but that they be united in it with others of like mind, and if possible with the whole human race. (6:135 – my emphasis)

Kant is conceiving of the kingdom of ends in terms of a divine state, and the principle of right is the principle that governs such a state. The principle of virtue tells us what creatures like ourselves need to do to be worthy of being members of such a state.¹⁹¹ In this chapter I will attempt to justify these claims.

¹⁹¹ In fact I think things are slightly more complicated than this, for the doctrine of right has both a pure and a impure part. The pure part of the doctrine of right examines the general principle of right. This is a principle that would govern an ideal civil condition. The impure part examines how we can attempt to make this world we live in closer to the ideal civil condition. The doctrine of virtue, then, has to do with what we have to do to become worthy of becoming members of such a state ourselves. The impure part of the doctrine of right has to do with how we can go about uniting the whole human race into such a state. That is, with what is involved in attempting to bring our actual political condition closer to the ideal. A virtuous man will strive for this. So we have a particular duty to try to improve the political situation here and now. So the impure part of the doctrine of right is subordinated to the doctrine of virtue; It only has to do with morality because we have a duty to try and do this.

(9b) – The Ontology of Property

Two types of interaction involving property

Recognizing that Kant conceives of the ideal of a civil condition as the idea of a real community, derived from the category of community, helps us understand Kant's account of property. In particular, it helps us to understand Kant's account of interactions between individuals involving property. At first sight it looks as if individuals can interact in two ways with regard to property: they can assert rights against one another, and property rights can be transferred from one individual to another. Ultimately, however, both of these types of property interaction are to be analyzed in similar terms. In the course of his discussion of property, then, Kant seems to describe two types of interaction, both of which are to be understood as conforming to the model of interaction Kant developed in his metaphysical work.

First, the mere fact of ownership is the basis of a form of interaction. Asserting a right against another person is to act upon her intelligibly. The fact that I own a particular object allows me to act upon you, for it permits me to legitimately demand that you desist from using it. Kant argues that property rights, and hence ownership as opposed to mere possession, are only possible in a civil condition. In other words, Kant believes that ownership is only possible in a community governed by laws. He makes it clear that when he talks of the civil condition he is not talking about some actually existing political community, but to an ideal of a community governed by laws. These laws are the basis of

legitimate rights and they have to be given (in the sense of being consented to) by each member of the community. Every time one claims a right and demands that it be recognized and respected by others, one is implicitly asserting the existence and validity of such a system of laws. It is only the existence of an (ideal) system of juridical laws, consented to by each member of the community, that allows one individual to act upon another by asserting her right against the other.¹⁹²

Second, when two individuals engage in a property transaction, they interact. In everyday language we talk of the ‘transferral’ of property rights, and this language seems to rest upon a crude influctionist conception of interaction, for it suggests that a property right is the sort of thing that can ‘flow’ from one individual into another. When I buy a painting from you it looks like a right that used to be yours becomes mine, as if a right is the sort of thing that can be transferred. If, however, like Kant, we reject the coherency of the crude influctionist conception of interaction we have to reconceptualize what is involved in such a ‘transferral’; we must be able to cash out what is involved in the ‘transferral’ in non-influctionist terms. One must read Kant's account of the ‘transferral’ of property with his rejection of a crude influctionist model of interaction in mind. Kant's alternative account of property transfers is based upon his account of action and interaction developed in his metaphysical works. He will argue that the ‘transferral’ of property rights is only possible if both parties constitute a community (or common will) and the ‘giver’ withdraws resistance to a claim that the recipient is actively asserting.

¹⁹² Although Kant argues that human positive laws are legitimate if they *could* be consented to, in an ideal community the members of the community must *actually* will the laws. When Kant claims that the criterion for the legitimacy of human laws is *possible* consent, what he means is that we *would* consent if we were fully rational.

Kant stresses the importance of simultaneity in property transactions. He rejects the idea that in the transferal of property there is first an *offer* followed by an acceptance of the *offer*. Instead Kant argues that the offer and acceptance must be simultaneous.¹⁹³ In arguing that offer and acceptance must be ‘simultaneous’ Kant is arguing that at the moment of ‘transferal’ the two parties must have a common [*Gemeinsam*] will.¹⁹⁴ If a transferal of property is to occur the giver and recipient must consent to a common set of laws governing property rights, and the recipient must actively assert a claim to the property while the giver must withdraw her resistance to this claim by simultaneously alienating her right.

Kant maintains, then, that in all property exchanges the giver merely withdraws resistance to a claim of the receiver. This suggests that, in some sense, Kant believes that giving is impossible, for the most we can do is allow others to take; we cannot give to others in the sense of filling them with something. This may seem problematic if we believe that ethics demand that we truly give to others. And Kant himself, as we shall see, believes that ethics, as opposed to the doctrine of right, demands that we give. Kant, however, does offer an account of what is involved in real giving. He believes that sometimes we have a (legal) duty to withdraw our resistance to the claim of others to an object. In such cases we are not really giving. We can, however, act beneficently if,

¹⁹³ Thus Kant explains that in a legitimate property exchange, “both acts, promise and acceptance, are represented not as following upon one another but. . . as proceeding from a single common will (this is expressed by the word **simultaneously**.)” (6:273 – my emphasis).

¹⁹⁴ These claims should be read in the context of his theoretical analysis of simultaneity, for in the *Third Analogy of the Critique of Pure Reason*, Kant argues that the category of community is the basis of judgments of simultaneity.

without having a duty to do so, we withdraw our rightful claim to the use of an object, and in so doing allow the claim of another to become a right,

Ownership

Kant would argue that many people's political beliefs are based upon a fundamental misunderstanding of what is involved in the idea of property. Because we own 'things' it seems natural to assume that ownership is to be understood as a relationship between an individual and an object. This is a fundamentally misguided way of conceptualizing ownership. For it does not explain what it is for someone to actually own something. To *own* something is to have a (legitimate) right to it. And to have a legitimate right to something is not to be understood in terms of the relationship between an individual and a thing owned, but instead is to be understood in terms of the 'owner's' relation to other agents. To claim a right is to claim that others should recognize your possession and not interfere with your use of the object. It is to claim that others should not resist your use, and Kant believes that such a claim can be made against others who consent to the same set of property laws.

Before we can explain what is involved in the transferal of property, we must understand what it is to actually possess property. Kant distinguishes between empirical possession and intelligible possession. To have a property right ultimately involves an intelligible

relationship.¹⁹⁵ Such intelligible rights, however, are only possible in the civil condition, thus Kant claims that, “only in a civil condition can something external be mine or yours” (6:256).¹⁹⁶

Kant conceives of the (idea of a) civil condition in terms of the idea of community. It is the existence of laws (governing property for example) which make us members of a community. It is these laws which ‘bind us’ together and make the community a community. The community is only a community if the laws are given (or consented to) by all members of the community. This, Kant believes is the case in an (ideal) civil condition. Each member of the community consents to the laws of the community and it is the existence of these laws that makes property rights possible. So, to begin with, the transferal of property assumes the existence of laws of property. In the *Metaphysics of Morals* Kant uses extremely Rousseauian language; he argues that,

the rational title of acquisition can lie only in the idea of a will of all united a priori (necessarily to be united), which is here tacitly assumed as a necessary condition (*conditio sine qua non*); for a unilateral will cannot put others under an obligation they would not otherwise have. – But the condition in which the will of all is actually united for giving law is the civil condition. (6:264)

¹⁹⁵ Thus Kant talks of “intelligible possession (*possessio noumenon*)”, and explains that property relations are “purely intellectual” (6:273).

¹⁹⁶ As the idea of a civil condition is an ideal that cannot be realized naturally (as it requires a belief in divine assistance) a consequence of this position is that in actually existing necessarily imperfect human communities property is also an ideal. Kant, however, believes that in such a communities we do have an (ethical) duty to respect the positive property rights that actually exist. His reasons for this are complicated, and his arguments not totally convincing. There seems no compelling reason why someone who buys into Kant's ontology of property should respect property rights in a society where many individuals do not have their needs met. Indeed, as we shall see, Kant himself believes that in such a society those in need have a *right* to demand assistance from the rich.

A few pages later Kant suggests that the possibility of ‘acting’ on another is dependent upon the existence of these system of laws. He argues that:

My possession of another’s choice, in the sense of my capacity [*Vermogen*] to determine it by my own choice to a certain deed in accordance with laws of freedom (what is externally mine or yours with respect to the causality of another), is a right (of which I can have several against the same person or against others); but there is only a single sum (system of laws), contract right, in accordance with which I can be in this sort of possession. (6:271)

The language here is very similar to the language he uses to explain action in his metaphysics lectures. There he argued that the agent must have a capacity to “determining the active power” of the patient. Here he claims that to have a right is to possess “a capacity to determine the choice of another”. And he argues that an individual can only possess such a capacity if there is a system of juridical laws and others (a) recognize and (b) affirm these laws. These laws are not physical laws but moral laws, the existence of which depends upon them being freely taken up by each individual member of the community. Kant explains that “my capacity to determine another’s choice by my own choices” is called a right and that it is the existence of juridical laws that makes rights possible and, as a consequence, allows one individual to act upon (“determine the choice of”) another. Laws that assign rights are called juridical (or coercive) laws. Such laws make interaction possible because they are the basis of resistance between individuals. In his ethical writings Kant repeatedly stresses that relation between such laws and the notion of resistance. For example, in his ethics lectures he argues that,

The universal law of reason can alone be the determining ground of action, but this is the law of universal freedom; everyone has the right to promote this, even though he effects it by **resisting** the opposing freedom of another, in such a way that he seeks to prevent an **obstruction**, and thus

to further an intent. . . The other, however, **obstructs** the action by his freedom; the latter I can curtail and **offer resistance** to, insofar as this is in accordance with the laws of coercion; so *eo ipso* I must thereby **obstruct** universal freedom by the use of my own. From this it follows that. . . the right to coerce the other consists in **restricting** his use of freedom, insofar as it cannot co-exist with universal freedom according to universal law; and this is the right of coercion. . . // Since nobody can exercise a right to coerce, who has not obtained a right thereto from a higher ground, which consists, however, in one's own freedom and its congruence with the freedom of everyone according to universal law, it is clear that the right to coerce can only be derived from the Idea of law itself. // Within this universal moral law are comprehended both legal and ethical laws. . . (*Ethik Vigilantius*, 27:523 – my emphasis)

We should read such passages bearing in mind Kant's account of action in his metaphysical work, for he believes that all action should be understood in terms of the withdrawal of resistance. Here Kant argues that that the right to coerce “consists in” (legitimately) resisting the freedom of others, and that such a right (i.e. the possibility of resistance) can only be derived from the “idea of law itself”. In other words in this passage Kant is suggesting that it is juridical laws that make resistance, and hence interaction, possible.

Kant believes that such a system of laws is only possible in the civil condition, a condition in which, “the will of all is actually united for giving law” (6:264). Kant believes that we are duty bound to (attempt to) enter such a condition. It is only in the context of such a system of laws I can legitimately demand that another withdraws her resistance to my claim to an object, for without such system of laws there will be no resistance, but also no possession. For, as Kant claims, possession is intelligible and involves an “intellectual title, and,

this intellectual title is the basis of the proposition: “what I bring under my control in accordance with the laws of outer freedom and will to become mine becomes mine”. (6:274)

Transferral of property

Only if such a community (or civil condition) exists can an individual really own property and ‘transfer’ her property to another. In so doing individuals are able to act upon one another through mutual consent. The activity of the agent (giver) is the withdrawal of an impediment, the activity of the patient (receiver) is an active uptaking. In the transferral of property, then, a property right does not flow from the giver to the receiver. Rather, in the context of a commonly willed set of property laws, one party renounces a right while the other party simultaneously actively uptakes the right. Kant is very careful to make it clear that in the transferral of property there has to be more than merely the ‘abandoning’ or ‘renouncing’ of a right by the giver, and Kant’s reason for stressing this is his commitment to the principle of active inherence. For the receiver to really possess a right she has to be the active ground of the right. Thus Kant explains that transferral of property,

is only possible [and I suggest that by ‘possible’ here Kant means ‘conceivable’] through a common will by means of which the object is always under the control of one or the other, since as one gives up his share in the common undertaking [*Gemeinschaft*] the object becomes the other’s through his acceptance of it (and so by a positive act of choice.) Transfer of the property of one to another is alienation. An act of the united choice of two persons by which anything at all that belongs to one passes to the other is a contract. (6:271)

Just as, in general, a determination can only belong to a substance if the substance is the active ground of the determination, property can only belong to an individual if the individual is the active ground of the right. Acquiring a right to something is not something that can occur passively, instead one must actively asserting a claim, even in the case of receiving a gift. In an act of exchange, then, it is not as if the donor actively gives and the recipient passively receives. Instead, the receiver must be actively asserting a claim to an object and the donor merely withdraws her (legitimate) claim to it, withdrawing resistance to the recipient's claim. This is why Kant stresses that the recipient must accept the property "by a positive act of choice".

We can now see how we can explain Kant's account of the transferal of property in terms of the language of his account of action in his metaphysics lectures. In chapter six I explained that (1) an object at rest has a mere **capacity** for motion; A ball being spun around a fixed point is moving in a circle. However at every particular moment it is striving to move in a straight line along a tangent; at every particular moment it (2) has a **dead power** to move in a straight line. If the string is cut, the ball flies off along the tangent; the dead power has been allowed to (3) become a **living power**.

Applying this terminology to Kant's account of property, we can say that (1) the existence of a system of juridical laws makes possible the existence of a **capacity** to own an object, that (2) a mere claim on the part of an individual is like a **dead power**, and that (3) a right is like a **living power**. In alienating her right the giver withdraws resistance to the recipient's claim to be allowed legitimate use of an object, in so doing the giver allows a

dead power in the recipient (the claim) to become a living power (a right). If, however, there was no assertion of a claim by the recipient there would be nothing the giver to do to make an object in her possession the property of anyone else.

In the legal sense all commissive acts are really omissive

Such considerations lie behind Kant's claim that in the legal sense, strictly speaking all commissive acts are really omissive. Kant explains this in his lectures on ethics. He argues that,

It must be noted. . . that all coercive or juridical laws are prohibitive, and rely on the principle of not withholding from the other what belongs to him (*neminem laede*). (For the fact that both commissive and omissive actions are equally necessary for the performance of actions in a physical sense, makes no difference, since all commissive actions are omissive, *in sensu juris*.) (*Ethik Vigilantius*, 27:512)

Thus although on the phenomenal level an act, such a paying a debt, may appear to be an action on the part of the debtor¹⁹⁷, on the legal level all that is happening is that the debtor is allowing his creditor to use what is legally hers. In paying back the loan, the debtor has not really given his creditor anything. Kant believes that such an analysis can be applied to all property transactions and not merely to cases of repaying a debt. Thus he explains that,

I cannot give the other anything – he already has what belongs to him; . . . you are to leave the other his own, take nothing, abstain from all actions whereby you would detract from his rights. (27:512)

¹⁹⁷ “In terms of physical forces [i.e. on the phenomenal level], the payment of a debt is nothing else but an action *commissiva*” (27:512).

From the legal (juridical) perspective, then, Kant believes that it is impossible to really “give” anything to anyone. It is easy to see how such an analysis can be applied to repayments of debts and to exchanges of property, but what about acts of charity? In what sense can the giving of a gift be thought of as act of omission? To understand Kant’s attitude towards the gift, we need to look at his account of beneficence and his distinction between right and ethics. For Kant believes that while all obligations “founded on [the] principle [*juris*] are negative. Right, however, differs from ethics, which tells us to give” (27:512).

Kant’s ultimate opinion will be that even gift-giving must also be understood as omissive, and that such benevolent behavior is only possible in the context of a juridical community. For gift-giving is only possible within a political community in which there exists legally defined property rights. If such a community did not exist I would not be able to give anything (as nothing would be rightfully mine to give) and the recipient of the gift would not be able to receive anything (nothing could ever become rightfully hers). I can only ‘give’ the recipient something that the laws of this community says is rightfully mine, and something the laws say that the recipient can potentially have a right to. The recipient of the gift must assert his potential right and I must renounce my actual right. The difference between repaying a debt and giving a gift is not that debt-repaying is omissive whereas gift giving is commissive, but that in the case of repaying a debt the debtor has a (strict) duty to renounce his right whereas in the case of beneficence the gift-

giver does not have a duty (or at least does not have a strict duty) to renounce his right.¹⁹⁸ Beneficence, then, is only possible in the context of a civil condition. Thus, although we have an ethical duty to be beneficent, Kant believes that it is only possible to exercise such a duty in the context of a civil condition governed by juridical laws. Thus he claims that, “If all men were willing to act from benevolence merely, there would be no ‘mine’ and ‘thine’ at all, and the world would be a stage, not of reason, but of inclination” (*Ethik Collins*, 27:416).¹⁹⁹

(9c) – The Doctrine of Right and the Doctrine of Virtue

The principium juris and principium ethicum

The *Metaphysics of Morals* is divided into a *Doctrine of Virtue* and a *Doctrine of Right*. And this may suggest that Kant postulates two distinct kinds of obligation: juridical obligations and ethical obligations. Further support for such a distinction is provided by the fact that Kant often distinguishes between ethical duties and legal duties (or duties of right).²⁰⁰ And, especially in his lectures, he makes a distinction between juridical laws and ethical laws. This suggests that we have two distinct ideas of community: the idea of

¹⁹⁸ Kant is not always consistent about this. For example, Vigilantius writes in his lecture notes that, “professor Kant maintains that in the state, a poor man has gained the right to demand support from the wealthy; for if it were left to his unrestricted choice, it would be perfectly open to him to earn so much for himself that he could make provisions for hard times” (*Ethik Vigilantius*, 27:540). This suggests that the poor have a right to be supported by the rich in a state, and so that in a state the rich have a coercive duty to show beneficence towards the poor.

¹⁹⁹ The only gift I can really give the other is the gift of consenting to a community in which such rights are possible. In other words, the only gift I can really give is my own autonomy which is the basis of the political community.

²⁰⁰ Explaining in his ethics lectures that, “law is the totality of all our compulsory duties (*leges strictae*). Ethics, the totality of all non-compulsory duties” (*Ethik Mrongovius*, 29:620).

a political community governed by juridical laws and the idea of an ethical community governed by ethical/moral laws. Although there are passages that support such an interpretation, ultimately Kant does not think that the difference between the juridical and the ethical has to do with content, for he believes that the idea of a purely ethical (as opposed to juridical) community is unintelligible.

In saying this I disagree with commentators such as Rosen (1993), who argues that “juridical duties are a proper subset of ethical duties” (p.88); I also reject the position that ‘ethical duties’ and ‘juridical duties’ should be understood as two species of a single genus. Instead, I argue that juridical duties and ethical duties are radically different types of things; the word ‘duty’ is being used in a different sense in these two phrases.²⁰¹ Juridical duties are merely the correlative of juridical rights. To have a juridical duty is merely to recognize that another individual has a legitimate right. Juridical duties are not to be understood in terms of necessitation. There is nothing contradictory in the idea of a holy being (that is a being not subject to necessitation) recognizing and respecting the rights of others. The notion of juridical duty has nothing to do with motivation. Ethical duty, on the other hand, essentially has to do with motivation. To do ones duty, in the ethical sense, is to be motivated in a certain way. Ethical obligations are only possible for beings such as ourselves whose sensuous nature means that our inclinations sometimes conflict with what we recognize as right. To be juridically obligated is not defined in terms of being (ethically) necessitated to do something, although for sensuous beings such as ourselves juridical duties are (ethical) obligations that often conflict with our

²⁰¹ In the *Metaphysics of Morals* Kant himself at times seems to try to draw a distinction between (juridical) duties and (ethical) obligation. See, for example, 6:383.

sensuous needs. Although Kant often talks of ethical duties in the plural, ultimately there is only one ethical duty, the categorical imperative, and particular ethical duties are merely aspects of this single duty. There is, however, a real plurality of juridical duties, for every right asserted implies a corresponding duty in others. Our ethical duty is to strive to be a (potential) member of a kingdom of ends. This pure idea is our ethical ideal, and ethical duties are an aspect of the schematization of this ideal.²⁰² In other words, Kant's taxonomy of ethical duties (plural) is an attempt to describe how the ideal of being a member of a kingdom of ends imposes itself on beings with sensuous natures such as we are. The idea of a kingdom of ends is the idea of a community of individuals that have rights and respect each others rights. As such it is the idea of a community of individuals who have juridical duties, although, as we have seen, it is not the idea of a community whose members are ethically obligated.

To understand Kant's account of the relationship between the ethical and the juridical we need to look at how he explains the distinction between the *principium juris* and the *principum ethicum*. Vigilantius writes in his lecture notes that,

professor Kant locates the supreme *principum juris* **in the limitation of anyone's freedom, through reason, to the condition that the freedom of each concur with the freedom of everyone, according to universal law**. He deduces from this, as a *corollarium*, the authorization to resist, or a right of coercion, insofar as the freedom of the other's action would violate the supreme principle of right, i.e., that the other's freedom would

²⁰² I am suggesting here that pure ideas, like categories, can be schematized, although ideas can only be schematized practically, whereas categories can be schematized theoretically. Categories can be applied *a priori* to intuition because our faculty of intuition has an *a priori* form: space-time. A schematized category is the rule for the production of an image in pure intuition. Ideas can be applied *a priori* to desire (or willing) because our faculty of desire has an *a priori* form: to act upon maxims. A schematized idea is a rule for the production of a maxim.

infringe upon your freedom, which coincides with the freedom of everyone according to universal law. (*Ethik Vigilantius*, 27:539)

The principle of right, then, demands that we respect the legitimate rights of others, and that they respect our rights. Such rights can only exist in the context of a system of juridical laws. The ideal of a civil community is the idea of a community that is based upon juridical, or *coercive*, laws. This is not to say that it is the idea (à la Hobbes) of a community in which there must be some supreme coercive power. Rather, civil or juridical laws are coercive in the sense that they specify rights, and these rights “obligate others to a duty”. In claiming this I am disagreeing with a common, if not the standard, account of the essential distinction between ethical and juridical laws. According to this approach, exemplified by Rosen (1993), who has written one of the few recent book-length studies on Kant's theory of justice, is that, “what essentially distinguishes a duty of justice from an ethical duty is that the former are enforceable but the latter are not” (p.110).²⁰³ While it is true that Kant believes that ethical duties are not enforceable, enforceability cannot be the distinguishing feature of distinctly juridical duties (and rights) for Kant himself recognizes a class of juridical rights that are not enforceable, namely what he calls rights of equity. Rosen recognizes the problem this causes for his account, but the only solution he can provide is to suggest that Kant must be mistaken to include duties of equity amongst juridical duties. Rosen explains the problem in the following terms:

Assuming that enforceable duties must correspond to enforceable rights, and that unenforceable duties must likewise correspond to enforceable rights, then all juridical rights should be enforceable, and hence narrow,

²⁰³ On this interpretation, juridical laws are coercive in that they are, or at least could be, enforced by a coercive power.

whereas all ethical rights should be unenforceable, and hence wide. // This conclusion is unavoidable. But it makes wide, unenforceable juridical rights impossible. (p.110)

However, as Rosen acknowledges, the fact that Kant seems to draw the conclusion that unenforceable juridical rights are, by definition, impossible, this does “not prevent him from asserting their existence in the *Rechtslehre*” (ibid.). Kant then seems to contradict himself, and Rosen ultimately believes that the most charitable interpretation of Kant’s position is that he actually does contradict himself.²⁰⁴

Rosen argues that Kant seems to allow for the possibility of unenforceable juridical rights in the *Doctrine of Right*, and what he has in mind are Kant’s claims about *equity* and the, so called, *right of necessity*. Thus Kant argues that,

An authorization to use coercion is connected with any right in the narrow sense (*ius strictum*). But people also think of a right in a wider sense (*ius latium*), in which there is no law by which an authorization to use coercion can be determined. There are two such true or alleged rights, *equity* and the *right of necessity*. (6:233-4)

Although it can be argued that Kant rejects, as spurious, the so called right of necessity²⁰⁵, it seems clear that he regards rights of equity as genuine rights, albeit rights that cannot be enforced. Thus Kant continues by arguing that equity “admits right without coercion” (6:234) and he recognizes that this seems to contradict his own position, but he believes that the conflict is merely apparent and that the “ambiguity really arises from the

²⁰⁴ From Rosen’s perspective, one could avoid the contradiction either by arguing that equity rights are enforceable, or by arguing that they are not really juridical rights.

²⁰⁵ Although I’m not sure he does actually regard the ‘right of necessity’ as spurious. He does argue that positive law cannot accept the right of necessity, but I believe that there are reasons to think that from the perspective of natural law we have a right to what we need. A discussion of this topic is beyond the scope of this dissertation.

fact that there are cases in which a right is in question but for which no judge can be appointed to render a decision” (6:234).

Kant, then, seems to suggest that equity provides us with examples of genuinely juridical rights that, in principle, cannot be enforced because it is impossible to appoint a judge who could legitimately enforce them. If Kant is serious about this commitment, then enforceability cannot be the defining characteristic of juridical laws and duties. I believe that Kant is serious here, and, as an alternative to the position that the distinguishing feature of juridical laws is that they are enforceable, I suggest that the essential difference between justice (*Recht*) and ethics is that justice has to do with rights, whereas ethics does not. In other words I agree with Fletcher (1987), who argues that in the strict sense **there are no ethical rights**. On this interpretation, the defining characteristic of a juridical (as opposed to an ethical) law is that a juridical law assigns rights. A right gives one a *moral* claim against another. Such claims are recognizable by reason, in the sense that they can be recognized by the court of conscience as legitimate, but they are not, however, necessarily enforceable by an externally appointed judge.

There are textual problems for this interpretation, for there are passages in which Kant seems to explicitly define the defining characteristic in terms of enforceability. My general strategy for dealing with such passages is to argue that in such passages Kant is giving an account of the defining characteristic of *positive* juridical law, and not of juridical law as such. In other words, I suggest that Kant believes that, although it is a

defining feature of human justice (and positive juridical laws) that rights must be enforceable by an (external) judge, this is not a defining feature of justice as such.

On my interpretation, then, a juridical law is a law that specifies rights, and as Kant explains in his ethics lectures,

the right to coerce the other consists in restricting his use of freedom, insofar as it cannot co-exist with universal freedom according to universal law; and this is the right of coercion. (*Ethik Vigilantius*, 7:521)

All rights are coercive because to claim a legal right implies some duties in others. For me to claim a right is to demand that others respect my right and not interfere with (or resist) my use of an object. The only basis for this demand can be an appeal to a common set of shared juridical laws. These laws must be such that they impose duties upon us as a result of granting legitimate rights to others. For each legitimate right granted implies a duty in others to respect the right.²⁰⁶ Juridical duties, then, are duties that are based upon the (asserted) rights of others. Juridical laws specify in what situations I may legitimately assert a right, and if I choose to assert such a right I create a juridical duty in others. This is one reason why such laws can serve as the basis of real interaction, for through asserting a (legitimate) right to an object one produces an intelligible change in others.

Kant contrasts the *principium juris* with the *principum ethicum*. The principle of ethics has both a formal and a material aspect. Formally, the principle of ethics is to:

²⁰⁶ This is why only juridical duties are enforceable through coercion. For a law is only enforceable through coercion if it implies specific duties in others.

Act according to law for the law's sake, or do your duty from duty. Act, that is, not only according to the law's imperative, but perform the act also, merely because the motive of the action is the law itself. (*Ethik Vigilantius*, 27:541)

This formal principle allows us to distinguish (conceptually) between the legality and the morality of an action. An action may be in conformity to the law, but will not be a moral or virtuous action if it is done for the wrong reasons. Only an action performed by someone who has a principled character can be moral as well as legally correct. Kant, however, argues that in addition to this purely formal *principum ethicum*, there is a *principium ethices*, “which is material, since it itself determines the action that is to be done” (*ibid.*). This material principle is material in the sense that it allows us to distinguish between the ethical and the juridical not merely in terms of the motivation of the agent, but also in terms of content of the action. This material principle, then, seems to imply a distinct set of purely ethical (as opposed to juridical) duties. In his lectures, Kant explains that this principle runs: “Act so towards other men, that you *can will* that the maxim of your action might become a universal law” (*ibid.*), and he adds:

Here, then, the object is not universal freedom, but will in relation to the universal will. The universal will consists in the universal end of all men, and is called love for others, the principle of well-wishing, directed to the universal end of happiness. (*ibid.*)

Ethically, then, we have a duty to care for the well-being and happiness of all men. Kant calls this the duty of *beneficence*.²⁰⁷ The juridical principle, then, is merely negative, it merely tells us that we are to “leave the other his own, take nothing, [and] abstain from

²⁰⁷ In the *Metaphysics of Morals* Kant explains that the duty of beneficence is not a duty to love others in the sense of having a certain feeling towards them (a duty to have a feeling is, Kant believes, “absurd”), rather it is a duty to love others *practically*: “To do good to other human beings insofar as we can is a duty, whether one loves them or not” (6:402).

all actions whereby you would detract from his rights” (27:512). This principle tells us to renounce a claim to an object only if another individual has a legitimate right to it. The (material) ethical principle tells us to renounce our claims to some objects even if another individual does not have a legitimate claim to them. As we shall see, however, Kant believes that all such ethical duties (of benevolence) can, or at least should, be reduced to juridical duties. The reason this is possible is that from the perspective of natural law individuals have a right to what they need, although no system of human, positive laws can accept such rights. Although the right of necessity cannot be incorporated into a system of positive laws and be enforced by an external court, it is recognized by the court of conscience as a genuine right. The person in need cannot, Kant believes, appeal to a public court to alleviate his need. He can, however, appeal to the court of conscience, and this court can recognize that he has a genuine right, although a right that cannot be recognized, in principle, by positive law. Although, from the perspective of positive law, acting benevolently towards others is not to be understood in terms of respecting their rights, from the perspective of natural law, the law that governs the court of conscience, it is to be understood in terms of the recognition of rights.

(9d) The Idea of a Kingdom of Ends is the Idea of a Political State and Not the Idea of a ‘Community of Friends’.

Introduction – disagreements with Korsgaard and Hart.

When I first read Kant I believed that Kant distinguished between the idea of an ethical community and that of a juridical (or political) community and I thought that the idea of

an ethical community was more basic, in the sense that ideally the juridical community would wither away and that, for Kant, an ideal state would be governed by purely ethical relations. In other words, I used to think that Kant's moral ideal was something like a community of friends. In her influential book, *Creating the Kingdom of Ends*, Korsgaard (1996) seems to be conceiving of the kingdom of ends in such terms, arguing that the idea of a kingdom of ends is analogous to a community of friends as opposed to a political state. I now believe that Kant's considered opinion is that the only conceivable type of community is a juridical or civil community, and that there are no distinct intelligible ethical rights or relations, although there are distinctly ethical duties. These purely ethical duties, however, are not intelligible, because they are not pure, for only impure sensuous beings with needs can have such duties, and so in our pure idea of a kingdom of ends there can be no notion of such ethical duties.

To be virtuous is to take a certain attitude towards this ideal of a juridical community (or what Kant calls the civil condition), and the ethical duties Kant enumerates in the *Doctrine of Virtue* are duties that only apply to virtuous beings and not holy beings. As our idea of the kingdom of ends is the idea of a community of holy beings we do not (and cannot) think of ethical duties (or laws) as existing within such a kingdom, for we can only have ethical duties towards creatures with needs, and our ideal of a kingdom of ends is the intelligible idea of a kingdom of holy beings. In a kingdom of ends, then, there would be justice, but there could be no beneficence. The reason for this is that in such a kingdom there would be no need for beneficence, for our idea of such a kingdom is of a society in which every individual would be doing their duty gladly, and this is only

possible if all of their needs are met. However, we cannot be beneficent to beings who have no unmet needs. This is not to say that Kant regards beneficence as ethically unimportant, it is merely to make it clear *why* it is important. Beneficence is a virtue, and a, if not the, central virtue. But as a virtue it is only important for beings like ourselves who can only strive to be members of a community of rational beings, it is not part of our notion of what such an ideal rational community would be like, for our ideal of a rational community is the idea of a community the members of which are fully rational and in whom there can be no tension between their rationality and their needs. Acting, and being treated, beneficently is only important for imperfect beings such as ourselves who are subject to (apparently) non-rational needs.

Although Kant believes that the only way we can think of a community is as one governed by juridical laws and that we cannot conceive of a community governed by purely ethical laws, he does believe that there are duties, the content of which is distinctly ethical. In claiming this I disagree with commentators such as Hart, who argues that Kant's distinction between justice and ethics can be explained solely in terms of motivation. Thus Hart (1984) argues that the distinction Kant draws between “legal rules” and “morals” can be explained purely in terms of the fact that

while legal rules only require “external” behavior and are indifferent to motives, intentions, or other “internal” accompaniments of conduct, morals on the other hand do not require any specific external actions but only a good will or proper intentions or motive. (p.252)

On such an approach the principle of right (the principle that governs legal rules) tell us *what* to do, whereas the principle of ethics merely tells us *how* to it. The principle of right

tells us to respect the rights of others in our actions; the principle of ethics, in contrast, tells merely that we should respect the rights of others out of duty (and not, say, merely out of fear of punishment). This account of the distinction, however, fails to account for the distinction Kant draws between the purely formal principle of ethics, and what he calls the material principle of ethics. The formal principle of ethics has to do with one's motive for obeying the law. The material principle, however, as we have seen, does require (specific) external actions that are not, and cannot, be mandated by juridical laws, for it demands that we act beneficently towards others. This duty of beneficence is, Kant believes, a purely ethical, rather than a juridical duty. The reason for this is because the duty of beneficence is not mandated by the rights of others. However, although we have a duty of beneficence, Kant believes that this duty is not intelligible, in the sense that it is a duty that would not, and could not, exist in an intelligible world or kingdom of ends.

In the remainder of this chapter I will begin by examining and rejecting Korsgaard's interpretation of the kingdom of ends as a non-political community, before moving on to discuss in more detail Kant's distinction between the ethical and political (juridical), paying particular attention to the role of beneficence in his ethics.

Korsgaard's Position

In chapter three we saw that Swedenborg conceives of heaven as an intelligible community governed by laws of love, or benevolence. Korsgaard seems to attribute a similar position to the mature Kant, arguing that he is conceiving of the kingdom of ends

as an “association created by love” rather than being based upon “the narrow relation of political justice”. In contrast, I have argued that the idea of a kingdom of ends should be understood as a juridical (or political) community, governed by laws of justice that allow individuals to assert rights which implicitly demand the respect of others. Although Kant believes that there are distinctly ethical duties he does not believe that the idea of a community or “association” based upon such duties is intelligible. Thus Kant would object to Korsgaard’s (1996) interpretation of his position.²⁰⁸ Korsgaard argues that,

Kant thinks that justice is reciprocal coercion under a general will²⁰⁹, made necessary by geographic and economic association. . . . When we share a territory we may have a dispute about rights. But I may enforce my rights against you only on the understanding that you may enforce your rights against me, and in this way we make a social contract and constitute ourselves a state. . . . Friendship is a free and uninstitutionalized form of justice, where the association is created by love rather than geographical necessity, and regulated by mutual respect rather than reciprocal coercion. // But it is not merely the narrow relation of political justice, but rather the moral relation generally, that friendship mirrors. (p.192)

Korsgaard continues by suggesting that the kingdom of ends should be regarded as an ethical community analogous to a community of friends, whose association is created by love rather than justice. Korsgaard, then, maintains that justice is merely a requirement for beings subject to geographical necessity. Her reading of Kant, then, is diametrically

²⁰⁸ Kant makes his attitude towards the idea of friendship clear in his lectures on ethics. He explains that friendship is an idea not an ideal. (It should be noted that here Kant is not using ‘idea’ in its usual sense of ‘an idea of pure reason’). And argues that, “so far as [a] maximum is a measure in regard to other, lesser qualities, such a measure is an idea; but in so far as it is a pattern for them, it is an ideal. If we now compare the affectionate inclination of people to one another, we find many degrees and proportions in regard to those who share out their love between themselves and others. . . . The greatest love I can have for another is to love him as myself, for I cannot love anybody more than that. . . . The idea of friendship enables us to measure friendship” (*Ethik Collins*, 27:423-4). Morality, in contrast, demands an **ideal**. Kant is insistent that the moral ideal is a pattern or archetype we can strive towards, it is not the idea of a **maximum**. The idea of a kingdom of ends, then, must be an ideal, and so cannot be the idea of a community of friends, for the idea of such a community is a maximum and not an ideal.

²⁰⁹ Up until this point I believe Korsgaard is right, although I’m not sure that we would agree about what is meant by ‘coercion’ in this context.

opposed to mine, for she maintains, in effect, that the laws of love are intelligible whereas laws of justice are not, because the existence of such laws is (empirically) conditioned. They are empirically conditioned, Korsgaard suggests, because they can only exist given the empirical fact of scarcity of land. Korsgaard's position implies that laws of justice, being dependent upon facts about geography, can only exist for (phenomenal) beings that are conceived of as essentially existing in space, for they are dependent upon irreducibly spatial features of the earth. In contrast, I argue that juridical duties are intelligible (we can understand them by appealing to our pure idea of community) whereas duties of benevolence are not intelligible, for we can only have such duties towards sensuous beings, towards individuals who have needs, and for whom ethics is a matter of necessitation.

What is probably motivating Korsgaard's reading here are Kant's remarks about possession of land in section 13 and onwards of the *Doctrine of Right*.²¹⁰ Here Kant suggests that if the surface of the earth "were an unbounded plane, people could be so dispersed on it that they would not come into any community with one another, and community would not then be a necessary result of their existence on the earth" (6:263). Here Kant seems to suggest that property relations, and hence the laws of justice, are merely the consequence of a contingent geographical fact, namely that the earth is finite. Elsewhere in the *Metaphysics of Morals*, however, he argues that we have a duty to enter into a political state governed by laws of justice, and this implies that he believes that laws of justice are not necessary because of some contingent empirical fact, but are

²¹⁰ Korsgaard herself points to 6:322 & 6:256 to justify her reading, but these passages make no mention of geography. She obviously has in mind Kant's remarks about possession of land.

necessitated by reason. Thus, for example, he argues that, “by the well-being of a state is understood. . . . that condition in which its constitution conforms most fully to principles of right [i.e., what Korsgaard calls “the narrow relation of political justice”]; **it is that condition which reason, by a categorical imperative, makes it obligatory to strive after**” (6:318 – my emphasis). This claim, which is clearly something Kant is committed to, suggests that laws of justice are not necessitated by contingent geographical facts, as Korsgaard suggests, but are demanded by reason. We have an ethical duty to enter into such a political state, and so the laws of justice are not necessitated by geographical facts, but by the categorical imperative.

Kant's remarks about land are merely an attempt to show what is involved in the possession of phenomenal object, and are not a pure part of his account of intelligible possession.²¹¹ In other words, in making his claims about geography, Kant is trying to explain how our idea of a juridical community, and the corresponding idea of intelligible possession, can (and must) be applied to the phenomenal world. The idea of property, however, is not merely applicable to possession of phenomenal, spatio-temporal objects; intelligible possession is to be understood in terms of possession of objects in general and is not limited to possession of objects considered phenomenally. Of course, we can have no positive contentful notion of a non-spatio-temporal object, but there is nothing contradictory in the concept of such an object, and so there is nothing contradictory in the

²¹¹ It should be pointed out that work needs to be done on the relationship between Kant's account of possession and his attempt to apply this to phenomenal possession. Any satisfactory reading of Kant's ethics must explain the relationship between our ideal of a civil condition and actually existing states and systems of laws. What is the relationship between the fact that ethics demands that we strive to be potential citizens in an ideal intelligible state, and the fact that we actually are members of (imperfect, phenomenal) states? A discussion of this issue is beyond the scope of this dissertation.

idea of possession of such (non-spatio-temporal) objects. The fact that land is the sort of thing that can be owned is, Kant believed, due to the contingent fact that the surface of the earth is finite. If the earth were not finite, there could, and would, be no (juridical) laws governing the possession of land. This, however, does not imply that there could be no laws of justice if the surface of the earth were infinite. It merely implies that there would be no juridical laws governing the possession of land, not that there could be no justice.

To conclude: morality is only a matter of virtue for beings like us who have a sensuous nature. As a consequence of this, duties of virtue, such as the duty of benevolence, are only duties for beings such as us and are not duties for what Kant calls rational beings as such. A community of holy beings cannot be conceived as subject to laws of benevolence or love. As Kant explains,

Vices like virtues remain always human, and the maximum of evil, and of good, in devil and angel is merely an unattainable ideal. . . we are indebted to it for the picture of heaven and hell. (*Ethik Vigilantius*, 27:691 – my emphasis)

The idea of a kingdom of ends, however, is also an ideal; it is not the idea of a perfect human community. It is, Kant argues at the end of the *Groundwork*, “the idea of a pure world of the understanding as a whole of all intelligences, to which we ourselves belong as rational beings (though on the other side we are also members of the world of sense)” (4:463). As such an ideal, our idea of the kingdom of ends has to be such that membership is not limited to beings such as ourselves. We must be able to think of all conceivable “intelligences” as potential members of such a kingdom. Thus our idea of a

kingdom of ends must be one that makes it possible for us to conceive of a kingdom of holy beings as a kingdom of ends. However, holy beings can not be thought of as subject to laws of benevolence, for duties of benevolence (and hence laws of benevolence) are duties of virtue, and as a consequence are only duties for beings such as us who experience morality as necessitation, so we cannot conceive of the kingdom of ends as essentially governed by laws of benevolence or love. We can, however, conceive of holy beings as governed by laws of justice.²¹² Indeed, Kant seems to suggest that insofar as we think of ourselves as autonomous and members of a kingdom of ends we must think of ourselves as holy beings. Thus, in the *Groundwork*, he argues that the “better person” is “conscious of a good will” that “constitutes the law for his evil will as a member of the world of sense” (4:454-5). For such a person, “the moral “*ought*” is then his own necessary “*will*” as a member of an intelligible world [=kingdom of ends], and is thought by him as “ought” only insofar as he regards himself at the same time as a member of the world of sense” (4:455). Our will, insofar as we think of ourselves as a member of the kingdom of ends, then, is not subject to necessitation, and so cannot be thought of as subject to duties of virtue. Kant makes essentially the same point a page earlier. Here he explains the possibility of categorical imperatives in the following terms:

Categorical imperatives are possible by this: that the idea of freedom [=autonomy] makes me a member of an intelligible world [=a pure world of the understanding = a kingdom of ends] and consequently, **if I were only this, all my actions would always be in conformity with the autonomy of the will** [i.e. I would be a holy being]²¹³; but since at the

²¹² The idea of a (finite) holy being is the idea of a being who cannot be thought of as *acting from duty*. Such a being is not necessitated by morality. We can think of such a being as having duties, in the sense of recognizing the rights of others. We cannot, however, think of such a being as having duties of virtue, for such duties are only possible for beings that are necessitated..

²¹³ As Kant explains: “A will whose maxims necessarily harmonize with the laws of autonomy is a holy, absolutely good will.” (4:439). He adds that, The dependence upon the principle of autonomy of the will

same time I intuit myself as a member of the world of sense, they ought to be in conformity with it. (4:454 – my bolding)

Here Kant makes it clear that insofar as I think of myself as a member of a kingdom of ends (an intelligible world) I think of myself as holy, for I think of myself as always and necessarily acting autonomously. The idea of virtue, and hence all duties of virtue, are only possible because I intuit myself as a member of the world of sense. A being that is not thought of as a member of the world of sense cannot be thought of as governed by laws of virtue. This is why Kant maintains that,

All moralists and teachers should . . . see to it that, so far as possible they represent acts of benevolence to be acts of obligation, and reduce them to a matter of right. (27:417)

The reason for this is because the moralist should present the ethical ideal in all of its purity, and benevolence is only an ethical matter for beings like us who have an impure dual nature. This is not to say that the ethical human being should not be benevolent, but benevolence is a result of striving to be a potential member of an ideal juridical community. In the imperfect political communities we live in, communities that are striving to instantiate the ideal of a civil condition, but whose laws are necessarily imperfect, we, as individuals, have an ethical duty to act benevolently. In an ideal civil condition, however, there would be no need of benevolence, or private acts of charity, for in such a society there would be a just distribution of property, indeed, for Kant, an ideal

that is not absolutely good (moral necessitation) is obligation. This, accordingly, cannot be attributed to a holy being” (*ibid.*).

society is one in which everyone has their needs met.²¹⁴ Benevolence is, and can, only be necessary in an imperfect political society, and we should not understand our acts of benevolence as supererogatory, but as acts that would be legally required if the society were really perfect, for in such a society everyone would have a right to what they needed and such rights would always be respected. Giving to those in need is not required by our imperfect laws governing property, but is an attempt to bring our societies closer to an ideal civil condition. Acting out of charity is not to respect an empirical right, but should be regarded as an attempt to bring our imperfect system of empirical rights closer to an ideal society governed solely by respect for rights. Such a society, however, could only exist if everyone has their needs met. Acting benevolently, then, involves respecting the rights people should have, not the rights they are actually assigned by our imperfect laws. The only basis we have for recognizing actual property rights are the rights assigned by human laws we live in. The need for benevolence is based upon the imperfections of human nature and the societies we live in and the laws that govern them. The need for benevolence, then, comes about as a result of the divergence between the laws that would govern an ideal state and actual human laws. A part of this divergence is contingent, and we recognize that most human laws could be improved. So we have a duty to attempt to improve the laws of the societies we live in. Another part of this divergence, however, is a necessary consequence of the fact that we are sensuous being, who have seemingly conflicting needs and who experience the world phenomenally. Kant himself gives a number of examples of how examples of benevolence can be represented as acts of

²¹⁴ Guyer (2000) makes a similar point in *Kantian Foundations for Liberalism*, when he argues that, “since there can be no rightful “unilateral acquisition” of property, there can be a rightful claim to property only within a system of “distributive justice”” (p.239).

obligation, and reduced to a matter of rights, and these examples make it clear that he regards the need for benevolence as a result of the imperfection of human societies and laws. Kant makes this clear in his treatment of equity.

At first glance, Kant's treatment of equity seems to be incompatible with his claim that (legitimate) rights necessarily imply duties in others. For he argues, in his lectures on ethics, that

Equity is a right, but one which gives no authority to compel the other. It is a right, but not a compulsive right. If anyone has worked for me, for example, for agreed payment, but has done more than I required, then he has, indeed, a right to demand payment for his extra work, but he cannot compel me to do it. (27:433)

The reason for this is that an external judge can only make judgments concerning the external actions of individuals and not their motivation. As a result of this our positive laws can only govern external actions and not internal choices. An ideal set of laws, however, would assign rights and govern the relations between individuals based upon their internal motivation and not upon their external behavior. There is no way, however, that purely human laws could ever satisfy this requirement. Kant makes this clear elsewhere when he argues that strict (positive) right can only govern explicit agreement and not implicit ones, although right itself governs both explicit and implicit agreements.

Thus Kant explains that,

The laws of right rest either on *jus strictum* (strict right), i.e. all the laws of coercion, or fairness, *aequitas*. The latter is a subtle concept, not yet sufficiently developed. It consists in the right to compel another, insofar as the latter is implicated in an undeveloped condition. The condition for coercion is therefore present only insofar and under such circumstances as have not been outwardly acknowledged, but which if they were so, would

establish the right to coerce; for example, the wage of a servant has been settled, but during his period of service the real worth of the currency is devalued, and we cannot defray his expenses with the amount in question (27:532-3)

Ideally our laws would regulate not merely explicit agreements, but also implicit ones. Kant, however, believes that it is impossible for an external (human) judge or set of positive laws to govern implicit agreements, for such laws can only govern what is external and hence what is explicit. This, however, is merely a limitation of human societies and not a limitation of justice itself. For, Kant believes, we are quite capable of recognizing that it would be unjust, according to the perspective of natural law, not to pay our servant enough to defray his expenses in times of inflation, for in hiring him we have implicitly agreed to pay him enough to live on. The judge of our conscience, unlike an external judge, is capable of recognizing implicit agreements as legally binding, and tells us that it would be unjust not to honor any such implicit contract. No statutory law can be drawn up, and no external judge appointed, who could justifiably enforce such agreements.

Bibliography

All references to Kant's writings, lectures and correspondence, except references to the *Critique of Pure Reason*, are given by volume and page number of the Akademie edition of *Kant's gesammelte Schriften* (Berlin, 1900-); the *Critique of Pure Reason* is cited by the standard A and B pagination of the first (1781) and second (1787) editions respectively. Unless otherwise stated, translations are from the Cambridge Edition of the Works of Immanuel Kant.

Ameriks (1992), Karl. "The Critique of Metaphysics: Kant and Traditional Ontology". The Cambridge Companion to Kant, ed. P. Guyer, 249-279.

Ameriks (2000), Karl. Kant's Theory of Mind – Second Edition. Oxford: Oxford University Press, 2000.

Beck (1966), Lewis White (ed.). Eighteenth Century Philosophy. New York: The Free Press, 1966.

Beiner & Booth (1993), Ronald & William (eds.). Kant and Political Philosophy. New Haven: Yale University Press, 1993.

Beiser (1992), Frederick C. "Kant's Intellectual Development: 1746-1781." The Cambridge Companion to Kant, ed. P. Guyer, 26-61.

Bellin & Ruhl (1985), Harvey f. & Darrell (eds.). Blake and Swedenborg. New York: Swedenborg Foundation, 1985.

Benz (2002), Ernst. Emanuel Swedenborg: Visionary Savant in the Age of Reason. West Chester: Swedenborg Foundation, 2002.

Brandt (1995), Reinhard. The Table of Judgments: Critique of Pure Reason A 67-76; B 92-101. Translated by Eric Watkins. Atascadero: Ridgeview Publishing Co., 1995.

Descartes (1984), René. Philosophical Writings, Volume II. Translated by Cottingham, Stoothoff & Murdoch. Cambridge University Press, 1984.

Di Giovanni (1992), George. "The First Twenty Years of Critique: The Spinoza Connection." The Cambridge Companion to Kant, ed. P. Guyer, 417-448.

Dole (1997), George. A Scientist Explores Spirit: A Biography of Emmanuel Swedenborg. West Chester: Swedenborg Foundation, 1997.

Fletcher (1987), George. "Law and Morality: A Kantian Perspective". Columbia Law Review 87, 533-57.

Förster (1998), Eckart. (Ed.) Kant's Transcendental Deductions. The Three Critiques and the Opus Postumum. Stanford: Stanford University Press, 1989.

- Fulbrook (1983)**, Mary. *Piety and Politics: Religion and the Rise of Absolutism in England, Württemberg and Prussia*. Cambridge: Cambridge University Press, 1983.
- Gregor (1993)**, Mary. "Kant on "Natural Rights"". In Beiner & Booth (1993).
- Gregor (1998)**, Mary (ed): Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. Cambridge, Oxford, 1998. In Ronald Beiner & William James Booth (eds.), *Kant and Political Philosophy*. New Haven: Yale University Press, 1993.
- Gawthrop (1993)**, Pietism and the Making of Eighteenth Century Prussia. New York: Cambridge University Press, 1993.
- Guyer (1977)**, Paul. "Review of W.H Walsh, *Kant's Criticism of Metaphysics (1975)*". *The Philosophical Review* 86 (1977), 264-270.
- Guyer (1992)**, Paul. (ed.) *The Cambridge Companion to Kant*. Cambridge: Cambridge University Press, 1992.
- Guyer (1993)**, Paul. *Kant and the Experience of Freedom: Essays on Aesthetics and Morality*. New York: Cambridge University Press, 1993.
- Guyer (1998)**, Paul (ed.). *The Groundwork of the Metaphysics of Morals. Critical Essays*. Lanham, 1998.
- Hart (1994)**, H.L.A. *The Concept of Law: Second Edition*. Oxford: Oxford University Press, 1994.
- Hartman (1974)**, Robert S. *Translator's Introduction to Kant's Logic*. New York: Dover, 1974.
- Herman (1993)**, Barbara. *The Practice of Moral Judgment*. Cambridge: Cambridge University Press, 1993
- Höffe (1994)**, Otfried. *Immanuel Kant*. Tr. Marshall Farrier. Albany: SUNY, 1994
- Kuehn (2001), Manfred. *Kant: A Biography*. Cambridge: Cambridge University Press, 2001.
- Horn (1997)**, Friedemann. *Schelling and Swedenborg: Mysticism and German Idealism*, trans. by George F. Dole. West Chester: Swedenborg Foundation, 1997.
- Hume (1978)**, David. *Treatise on Human Nature*. London: Routledge & Kegan Paul, 1978.
- Johnson (1996)**, Gregory. "Kant on Swedenborg in the Lectures on Metaphysics". *Studia Swedenborgiana*, Vol. 10, No. 1.
- Korsgaard (1996)**, Christine M. *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press, 1996.
- Kuehn (2001)**, Manfred. *Kant, a Biography*. Cambridge: Cambridge University Press, 2001.
- Laywine (1993)**, Alison. *Kant's Early Metaphysics and the Origins of Critical Philosophy*. Atascadero: Ridgeview Publishing Company, 1993.
- Leibniz (1989)**, G.W. *Philosophical Essays*, ed. and trans. by Roger Ariew and Daniel Garber. Indianapolis: Hackett, 1989.

- Leibniz (1991)**, G.W. *The Monadology: An Edition for Students*, ed. and trans. by Nicholas Rescher. London: Routledge, 1991.
- Leibniz (1996)**, G.W. *Disputatio*, trans. by Laurence B McCullough in *Leibniz on Individuals and Individuation: The Persistence of Pre-modern Ideas in Modern Philosophy*. Dordrecht: Kluwer, 1996.
- Leibniz (1997)**, G.W. *Leibniz' New System and Associated Texts*, ed. and trans. by R.S Woolhouse & Richard Franks. Oxford: Clarendon Press, 1997.
- Longuenesse (1998)**, Beatrice. *Kant and the Capacity to Judge : Sensibility and Discursivity in the Transcendental Analytic of the Critique of Pure Reason*. Princeton: Princeton University Press, 1998.
- McCullough (1996)**, Laurence B. *Individuals and Individuation: The Persistence of Pre-modern Ideas in Modern Philosophy*. Dordrecht: Kluwer, 1996.
- McDannell & Lang (1988)**, Colleen & Berhard. *Heaven: A History*. New Haven: Yale U.P., 1988.
- Neiman (1994)**, Susan. *The Unity of Reason: Rereading Kant*. New York: Oxford University Press, 1994.
- Nietzsche (1961)**, Friedrich. *Thus Spoke Zarathustra*, ed. and trans. by R.J. Hollingdale. London: Penguin Books, 1961.
- O'Neill (1989)**, Onora. *Constructions of Reason. Explorations of Kant's Practical Philosophy*. Cambridge: Cambridge University Press, 1989.
- Pozzo (1998)**, Riccardo. "Kant Within the Tradition of Modern Logic: The Role of the "Introduction: Idea of a Transcendental Logic"". *Review of Metaphysics* 206 (1998), 295-310.
- Raine (1985)**, Kathleen. "The Human Face of God". In Harvey F. Bellin & Darrell Ruhl (eds.), *Blake and Swedenborg*. New York: Swedenborg Foundation, 1985.
- Rawls (1993)**, John. "Themes in Kant's Moral Philosophy". In Ronald Beiner & William James Booth (eds.), *Kant and Political Philosophy*. New Haven: Yale University Press, 1993.
- Reich (1992)**, Klaus. *The Completeness of Kant's Table of Judgments*. Tr. Jane Kneller and Michael Losonsky. Foreword by Lewis White Beck. Stanford: Stanford University Press, 1992.
- Rossi (1991)**, Philip J. and Wreen, Michael. Editors. *Kant's Philosophy of Religion Reconsidered*. Bloomington: Indiana University Press, 1991.
- Sartre (2001)**, Jean-Paul. *Being and Nothingness*, ed. and trans. by Hazel E. Barnes. New York: Citadel Press, 2001.
- Schilpp (1998)**, Paul. *Kant's Pre-critical Ethics*. Bristol: Thoemmes Press, 1998.
- Schiller (1967)**, Friedrich. *On the Aesthetic Education of Man*, ed. and trans. by Elizabeth M. Wilkinson and L.A. Willoughby. New York: Oxford University Press, 1967.
- Schneewind (1990)**, J. B. *Moral Philosophy from Montaigne to Kant: An Anthology*. Vol 1. and Vol. 2. New York: Cambridge University Press, 1990.

- Schneewind (1992)**, J. B. "Autonomy, Obligation, and Virtue: An Overview of Kant's Moral Philosophy". *The Cambridge Companion to Kant*, ed. P. Guyer, 309-341.
- Schneewind (1998)**, J. B. *The Invention of Autonomy: A History of Modern Moral Philosophy*. New York: Cambridge University Press, 1998.
- Schönfeld (2000)**, Martin. *The Philosophy of the Young Kant*. New York: Oxford University Press, 2000.
- Schopenhauer (1970)**, Alfred. *Essays and Aphorisms*, ed. and trans. by R.J. Hollingdale. London: Penguin Books, 1970.
- Semler (1781)**, D.J.S. *Lebensbeschreibung von ihm selbst abgefasst*, 2 volumes. Halle, 1781-2.
- Solovyov (2001)**, Vladimir. Article on Swedenborg in *Brockhaus-Ephron Encyclopedia*. Translated by George Dole. *Studia Swedenborgiana*, Vol. 12, No. 2.
- Stoeffler (1973)**, F.E. *German Pietism during the Eighteenth Century*, Leiden: E.J. Brill, 1973.
- Sullivan (1989)**, Roger J. *Kant's Moral Theory*. Cambridge: Cambridge University Press, 1989.
- Swedenborg (1995)**, Emanuel. *Heaven and its Wonders and Hell: From Things Heard and Seen*, trans. by J.C. Ager. West Chester, Pennsylvania: Swedenborg Foundation, 1995.
- Swedenborg (1997)**, Emanuel. *Heavenly Secrets*, trans. by John Clowes. West Chester: Swedenborg Foundation, 1997.
- Watkins (1995a)**, Eric. "Kant's Theory of Physical Influx". *Archiv für Geschichte der Philosophie* 77 (1995) 285-324.
- Watkins (1995b)**, Eric. "The Development of Physical Influx in Early Eighteenth-Century Germany: Gottsched, Knutzen, and Crusius". *The Review of Metaphysics* 49 (1995) 295-339.
- Wood (1999)**, Allen W. *Kant's Ethical Thought*. New York: Cambridge University Press, 1999.
- Woolhouse & Franks (1997)**, R.S & Richard (eds.). *Leibniz' New System and Associated Texts*. Oxford: Claradon Press, 1997.