COMPULSION OR ATTRACTIVENESS – A FALSE DICHOTOMY? A KANTIAN PERSPECTIVE ON THE SOURCES OF MORAL MOTIVATION

MARIA SCHWARTZ

University of Augsburg

Abstract. The essay questions the dichotomy between 'push' and 'pull' motivation to act morally, asking for the motivational power of Kant's categorical imperative instead, its functionality as well as its sources. With reference to Christine Korsgaard it can be shown that personal integrity together with the notion of an ideal common world form one single source of motivation, grounded in exercising our autonomy. In a last step this outline of a kantian ethics of automony is related to the notion of God, whose role is illustrated in Kant's *Religion Within* in a surprising way.

Since G. E. M. Anscombe's *Modern Moral Philosophy* (1958) many if not most scholars in ethical philosophy, Linda Zagzebski amongst them, open a sharp dichotomy between two types of ethics (cf. Zagzebski 2004: xi f.):

(1) An ethics of duty (deontological), associated with a commanding law, duty (the 'ought'), obligation, compulsion and force, concerned with morally right or wrong acts.

(2) An ethics of value or virtue (teleological), associated with attractiveness, thick moral concepts and being a good or virtuous person.

However, taking a closer look at the ethics of Immanuel Kant, I think this dichotomy is wrong. Specifically, there is no such conflict as compulsion versus attractiveness once we look at the sources of moral motivation from a Kantian perspective. My contribution in this article

aims at pointing out two main and closely connected sources of moral motivation, which underlie Kantian ethics:

(a) Personal integrity and self-conception as a free human being.

(b) The notion of other people in this world as free human beings.

Both of these sources are related to compulsion and attraction alike. They are grounded in valuing oneself and others as free human persons, living in a common world.

In the following I will ask 1) for possible categories of factors motivating our actions, drawing a very simple picture. My questions are a) Who or what is motivating? and b) *How* does motivation work? In part 2) I move on to take a closer look at the motivational power of the Categorical Imperative. In part 3) Conception of the Self and 4) Kingdom of Ends, I rely mainly on Christine Korsgaard's *Sources of Normativity* (1996) to discuss personal integrity and the notion of an ideal common world as motivating factors.

In the final part 5) I attempt to relate this outline of an ethics of autonomy to the notion of God, asking if there is any place left for religious approaches at all and if so, what they would be.

I. MOTIVATING FACTORS – THE (VERY) SIMPLE PICTURE

Let us consider a basic situation of moral challenge:

> A person P is in desperate need of help. I am the only one present and able to help by doing X. So X is the moral deed in question.

Question 1: What or who is motivating me to do X?

Once we ask *what or who* possibly is motivating me to do X, there are several everyday-answers like:

- God
- People
- Me, myself (Reason? Emotion? Will?)
- Values
- Duty

However, the philosopher intervenes instantly, pointing to at least three entirely different categories of concepts all regarded as 'motivating factors' here. Assuming God is a person these are:

> a) persons/people, b) human faculties and c) abstract concepts like values or duties.

Question 2: How are these factors motivating?

Once we ask in a second step *how* these factors should be considered 'motivating', the mentioned dichotomy between attractiveness and compulsion arises. Yet there are significant differences in how well that dichotomy fits to each category. Let us again consider some everyday-answers people might give when questioned about their motivation to do X. Applied to the first category, motivated by persons (from a Christian perspective, including God), a first, encouraging answer might be:

> I love God and his commands, so I do X. Or: I like the person in need, so I do X.

We can regard this motivation as a 'pull-motivation', working via attractiveness. It is attractive to do X, so I act accordingly.
Another, more sobering type of answer might be:

> God (or people) will punish me if I don't do X.

We can regard this motivation as a 'push-motivation', working via force or compulsion. It is not attractive to do X. However, something very bad will happen if I refuse to help, thus I feel compelled to do X.

In the whole area of relationships, dealing with persons, distinguishing between the two different kinds of motivation seems just fine. And it makes even more sense in a theological frame like 'Divine Command' versus 'Divine Motivation' theory.

Moving on to the other two categories the distinction becomes questionable, mainly because certain types of motivation (like attraction) are tied to certain concepts (like values). It appears that inner-personal faculties like reason and emotion or concepts like duty and value are *as* ambivalent as persons with regard to their motivational powers. Each concept or faculty is capable of motivation via compulsion or attractiveness. For example, to look at everyday-language once again, consider these sentences:

> 'I will feel remorse when I don't do X.' – 'I like to be a helpful, virtuous person.' – 'I love to follow my duty.' – 'I feel compelled to realize a value, to carry it into effect.' – 'I *have* to change the world, no matter what!'

As the last two examples indicate, strong values might attract somebody in a way that makes it extremely difficult to see any difference between a strong attraction and a compulsion at all. Pull easily becomes push and vice versa.

I am questioning if the distinction between 'push' and 'pull', between 'compulsion' and 'attractiveness' is helpful at all. As a side remark, Zagzebski's theory is much more complicated and tries to bind these concepts together. She would most likely regard this picture as over-simplified (where I would agree). It might also look different once there is no clear dichotomy between reason and emotions in the first place[1]. However, this simple picture will be useful for the following parts where I am going to outline a theory of autonomy, especially the concept of duty ('Pflicht'), addressing the kind of motivating power in Immanuel Kant's ethics (part 2) and its sources (part 3 and 4).

II. THE MOTIVATIONAL POWER OF THE CATEGORICAL IMPERATIVE

How to describe the motivating power of the Categorical Imperative? Is it based on attractiveness or force? The answer I am trying to give is: Neither, because Kant has a wholly different account of motivating forces, the 'Triebfedern'. The moral law is regarded as the sole, objective ground of motivation. To become effective, however, it has to become the subjective 'Bestimmungsgrund'. How does this happen?
A very interesting remark in Kant's *Groundwork* tells us:

> ... [the human being] is subject *only to laws given by himself but still universal* ... if one thought of him only as subject to a law (whatever it may be), this law had to carry with it some interest by way of attraction or constraint, since it did not as a law arise from *his* will. (G 4:432-433)

Therefore, attraction or constraint ('Reiz oder Zwang') is *not* how the law given by one's own will works. Why is it not motivating in the way of a push-motivation, by force? Kant is pointing out here that we are used to relating the term 'law' instantly to force, since when there are laws, there are also institutions which care for these laws to be enforced. We speak of 'armed forces' or the 'force' of the circumstances which leaves us 'no choice'. 'Force' could also describe the power to make people do something they do not want to do, be it by means of pure violence or by manipulating them otherwise. The latter might be the crucial point.

There is not the slightest force of that kind in the demands of the Categorical Imperative (CI). To be obligated means to know you *should*

[1] Zagzebski discusses cognitive and affective aspects of emotion (cf. Zagzebski 2004: 68).

do something, just because you know it is good to do – especially when there is nobody around to *force* you to do it.

There are other quite interesting remarks in Kant's later works, mainly the *Metaphysics of Morals* and the *Religion within,* where he assumes any incentive ('Triebfeder') has to be willingly included into one's maxim:

> ... freedom of the power of choice has the characteristic, entirely peculiar to it, that it cannot be determined to action through any incentive *except so far as the human being has incorporated it into his maxim* ... (Rel. 6:23-24)

Even the Categorical Imperative itself has to be willingly endorsed, be adopted as one's highest maxim, in order to be put into effect. At least this is what the *Religion within* tells us. Thus, to speak of attractiveness or compulsion is equally wrong, because the underlying principle of adopting moral maxims is freedom. Starting from the adoption of maxims and down to the specific actions we perform, it is always our very own freedom that enables us to decide. It is our free choice to do X, even and specifically *if* there are strong external or internal forces standing against a moral decision.

In both the *Metaphysics of Morals* and the *Religion within* alike a decision is needed for whatever good motivating factor, even the moral law itself. Otherwise it will have no effect on one's maxims, and consequently on one's actions, at all. The rational endorsement of any factor whatsoever, apparently seems to neither have nor need any further source, because at that point it would no longer be a free decision.

In our initial example you might expect to get punished for helping the person in need by an unjust regime or P might be your worst enemy. Yet you are the only one who is able to help, knowing it would be a good thing to do. This 'knowing I should do X' amounts to the moral obligation standing against other inner and outer forces alike.

Assume in our example that you would just walk away and nobody will notice your refusal to help. Nothing will happen to you if you just decide differently. You might escape punishment, which is normally considered a good thing. Or you might experience some inner satisfaction of paying your enemy back, successfully suppressing any forthcoming feeling of remorse.

Thus, the rigid 'commands' of the Categorical Imperative are probably the only type of commands in this world which are *free* from force. To be obligated by the CI means to be motivated (in a way I will

discuss shortly), but still always be free to do otherwise. From an outside perspective, its demands seem to be powerless, un-founded, so to say. Gerd Haeffner, a teacher of mine and Phenomenologist, once spoke of the harmlessness and innocence ('Unschuldigkeit') of the demands of the Categorical Imperative. I never heard a more apt expression to describe its motivating power.

When it comes to explain acting on duty ('aus Pflicht'), it is often stated this means doing your duty just because you know it is duty. To do a good deed just *because* it is a good deed seems like an empty command, however, it fits to the notions of choice and freedom.

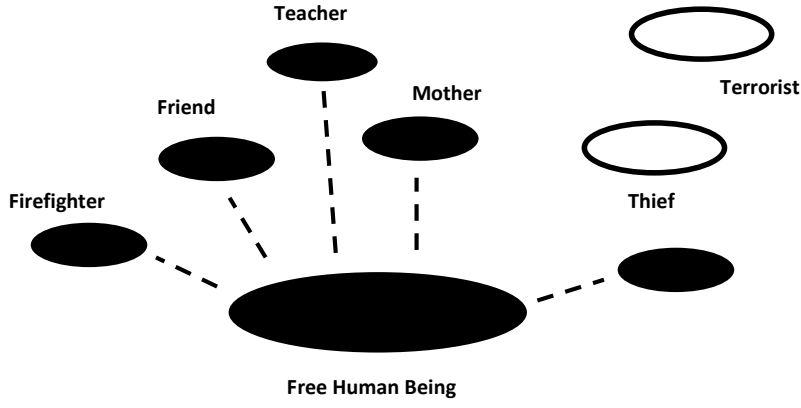
A brief side-remark regarding possible objections: Kant does speak of 'Nötigung' frequently (for example in G 4:413, G 4:434, MM 6:437), which could be translated as 'necessitation' or even 'force'. It is also difficult to see how the intellectual feeling of respect ('Achtung') fits into the motivational picture, which would be a separate topic. I think Kant is constantly struggling to hold on to the initial idea in the *Groundwork* that the CI is adopted due to rational, free decision alone, and nothing else.

Still, the standard interpretation that acting morally in Kantian ethics amounts to following duty for duty's sake might be a preliminary, outside view. Contemporary Kant scholars like Christine Korsgaard are asking further questions. We are surely not moral for *no* reason, just to be moral. The moral saint will rescue her enemy, even if she might endanger her own life doing so. In answering the normative question of *why* she would act in this way, there must be more to say (cf. Korsgaard 1996: 13, 17-18). Why should moral demands bother us at all if they were only suggesting 'doing duty because it is duty'? Is there really no further explanation? This question leads us to part 3), Korsgaard's motivational account.

III. PRACTICAL IDENTITY – OUR CONCEPTION OF THE SELF

Christine Korsgaard offered a very interesting answer to this question with regard to personal identity, combining the Kantian with a Platonic approach (cf. Korsgaard 1996). To understand her thesis in a nutshell, again considering our initial example: I cannot 'live with myself' any longer if I do not do X because my deepest identity consists in acting according to the CI. A human being's identity first and foremost consists in acting as a free person, in making her or his own free choices. The actual place where freedom fulfils itself is the moral realm.

Kant assumes there is a hierarchy of different classes of imperatives and accordingly of different maxims[2]. What Korsgaard additionally assumes is that we are acting in the light of a hierarchy of different practical identities or conceptions of ourselves[3], illustrated by the following graphic:



Identity as a free human being (in my graphic at the very bottom) is connected to the CI. In this light I choose all my different contingent social identities, which could also be regarded, in the Kantian terminology, as whole sets of maxims that we are trying to live up to.

Contingent identities are such as the identity of being a good friend, mother, teacher, or firefighter. All of those come with different duties. If a teacher cannot get up in time to hold his classes he will probably lose this identity sooner or later or at least not be regarded as a good teacher. The dotted lines in the graphic indicate all those identities which could be chosen in accordance with the Categorical Imperative. There are also some which can't be chosen without damaging my fundamental identity, like the one of a terrorist or thief. Both identities include using other people as a mere means to one's ends, impairing or even eliminating the freedom of others, so they must not be chosen. How it is even possible to choose an identity like that, willingly, which amounts

[2] An imperative suggests something I *should* do. As soon as I answer positively, deciding that this is something I really *want* to do, it becomes my maxim (cf. Schwartz 2006: 44-76).

[3] Cf. for the following Korsgaard 1996: 120-123.

to the abandonment of a fundamental part of what makes us human, remains an unsolved mystery in the *Religion within*, where Kant speaks of the radical evil (for example Rel. 6:32, 6:35, 6:43)[4]. Once chosen, the practical identities are a powerful source of motivation or, as Korsgaard states, the source of obligation:

> It is the conceptions of ourselves that are most important to us that give rise to unconditional obligations. For to violate them is to lose your integrity and so your identity, and to no longer be who you are. ... An obligation always takes the form of a reaction against a threat of a loss of identity. (Korsgaard 1996: 102)

Back to our initial question about the kind of motivation: If not the Categorical Imperative itself, are practical identities perhaps providing a push- or pull-motivation, based on attractiveness or force? Judging from the latter quote, Korsgaard rather seems to think it is a push-motivation, a negative one, grounded in the fear of damaging one's integrity.

However, I would like to add that if you are living up to an identity, for example, as a good friend, by choosing not to betray your friend you are strengthening your identity. On closer inspection it might be very attractive to not only preserve, but live up to the identities we value. The feeling of respect ('Achtung') also points in this direction. After doing X one could say: 'This might have been hard, but I'm proud I did X. It was the right thing to do.' It does not seem to be only the fear of losing one's identity, which might motivate some of our actions, but also the positive feeling of forming a character and living up to one's true self.

This ambivalence especially applies to the most fundamental identity, the one of being a free human being. An insight of spiritual literature (for example in the works of St. Ignatius of Loyola) is that the more you practice your freedom, the more free you are becoming. This bottom source of motivation, one's human identity, can neither be categorized as compelling nor attracting. It is about freedom, lying above or in between.

And yet this approach faces some common criticism. If my self-conception is the underlying source of moral motivation, wouldn't it just be my own selfishness motivating me, striving for preserving my

[4] The inclinations ('Neigungen'), can't be regarded as root of evil choices; these are only responsible for *akrasia*. As Kant writes, inclinations are opponents of principles in general ('Grundsätze überhaupt'), whether they are good or bad (cf. Rel. 6:59, Footnote). Once a person suffers from *akrasia,* she can't even rob a bank successfully. The free decision of the will is necessary condition for good and evil actions alike.

identity? I think that is only one side of the coin. One further step is left to answer this objection which I will outline in the following part 4.

IV. THE KINGDOM OF ENDS – BEING PART OF A COMMON WORLD

Taking a closer look at *what* the Categorical Imperative actually commands adds a whole new dimension to the picture. The CI is not concerned with specific advises like 'Do this' or 'Don't do that'. It is a formal principle. Thus, to understand the motivational force of the CI, the formulas and their application need to be discussed. To consider one of its most known formulas in the *Groundwork*:

> '... act only in accordance with that maxim through which you can at the same time will that it become a universal law.' (G 4:421)

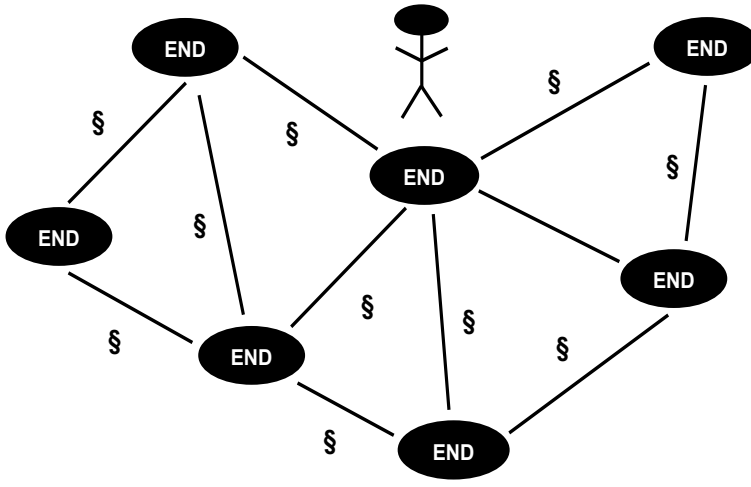
What does it mean to 'will that it become a universal law'? How are we to judge what should become a universal law and what not? To even try to do so, I have to take the view of other people into account, treating them as ends in themselves. The universal law is not exactly a random law I am giving *myself*, stuck in some sort of solipsism. To focus on autonomy ('Selbst-Gesetzgebung'), like we did so far, would be only one – and an incomplete – description, omitting the vital point. Once we take a closer look at the formula, it opens a new horizon towards other people: The horizon of all of us forming a common ideal world, a Kingdom of Ends.

The reason why many people regard the Categorical Imperative as implausible, not matching our everyday experience, is that this horizon is never fully conscious. Though we are acting in its light all the time, we are not doing so consciously. Nobody asks himself 'can I universalize the maxim underlying my action?', deriving it step by step from the CI as if in a mathematical operation. All of this is happening way faster, parallel to deliberation or rather, forming the background of our decision making.

If we want to apply the Categorical Imperative consciously, like we do in ethical theory, it requires nothing less than imagining a whole ideal world in which we all want to live. Becoming a law means that everybody is allowed to act like me, in a similar type of situation. It further means that I will grant all my reasons justifying my action to everybody who acts the same way, while I am affected this time. The notion of universal law requires taking other human beings seriously, which leads us instantly to the second formula, to treat them as an end, not as mere means:

> '*So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means.*' (G 4:429)

And once we combine both of these formulas, we arrive at the formula of the Kingdom of Ends. Like the following graphic shows, there must be a number of persons to whom the law applies (here pictured as dots):



The structure provided by the universal law-formula necessarily has to be 'filled' by people which are ends in themselves. Laws regulate the relations between people. Once I am doing something which affects another person, it always should be something which I also agree to be done to me, justified with same relevant reasons.

A side remark on the importance of reasons given in the maxims: For example, suicide is not regarded as a moral decision in the *Groundwork*. The reason for this verdict, however, is not the obvious one that I am getting rid of my humanity altogether (since then suicide would be prohibited under all conditions). It is rather the reason the agent himself provides for his decision: The suicidal person in the example is neither sick nor about to be executed. He simply doesn't expect the rest of his life to be more pleasant than unpleasant (cf. G 4:422). In the *Metaphysics of Morals* Kant questions if suicide shouldn't be allowed in other cases, for other reasons, like committing suicide in order to rescue his country

or to save others from a contagious illness (cf. MM 6:423f.). So the universalization of maxims includes a universalization of reasons. Not mere acts are assessed, but whole rich descriptions of actions, including circumstances and reasons, like they are already found in the examples of the *Groundwork*.

Applying the CI as a formal principle requires 'filling' it, considering all the values, emotions, the specific factual needs and interests of other people. Kant himself is not very much bothered with this empirical base, though. He is only unfolding the basic structure and theory.

What does the notion of the Kingdom of Ends add to our understanding of motivational sources? After pointing to one's own integrity in part 3, we now possibly notice an even deeper motivating factor, which stems from our humanity *as* being part of a common world full of relationships and other people. What does it mean that these people are not allowed to be used as a mere means to our personal aims? They have to agree to our actions, *in case* these are constraining their legitimate freedom.

Like we saw in part 2, however, there is not the slightest force making us take the view of other people into account which are affected by our actions. We are free to ignore the common world idea, anytime. Most people committing crimes know perfectly well the world wouldn't function anymore if everybody acted like them. And most even rely on others *not* doing the same, making exceptions for themselves, as Kant writes in G 4:424. The common characteristic of immoral actions is that nobody could rationally wish to live in a world where such courses of action become a law.

The advantage of the criteria of universalization is that it does not stay as opaque as an ethics of moral examples. Reading Aristotle can leave us puzzled and helpless when it comes to the question of what exactly is a moral thing to do. I simply do not know what the moral expert (a *phronimos*) would do if he were me. Like John Hare pointed out (cf. Hare 2005) most of us probably even more lack the empathy to be able to judge what God would do. To even identify moral examples like St. Francis of Assisi or Gandhi, which are suggested by Zagzebski (cf. Zagzebski 2004: 46), we need to have a notion on what exactly is good or bad in the first place. It requires a lot of imagination and empathy to see what a morally wise person would do in my situation. Answering the question in which kind of world we would like to live in, on the other hand, is much easier. Most likely this will be a world of peace, of respect and caring, no cheating, no stealing, no killing and so on.

To summarize: First, like outlined in this part, I do not think human autonomy in a Kantian framework has to do with solipsism or selfishness at all. The motivational power of the CI stems from a fundamental identity, which is inseparably connected to the world of all human beings around me.

Secondly, as shown in part 2, since all moral decisions are free decisions, both approaches, be it a push- or pull-motivation, are actually not applicable to the Kantian framework. Considering our example of doing X, it does not matter if it is attractive to help somebody – for example my best friend or my child who is in danger – or if strong forces of whatever kind are holding me back. It would be moral to just help, not because there is a strange value to be realized through that deed, but because there is another human being in need of help, which is valuable as a person.[5] Nobody wants to live in a world where people make it a law to ignore each other. If I would constantly demand others to treat me as an end in itself while at the same time excluding myself from this demand, treating everybody as a pure means to my ends, I would be acting immorally. This is also not how God created us or meant us to be. Finally this leads us to a few brief remarks on the relation of an ethics of autonomy and the religious dimension.

V. RELATING AN ETHICS OF AUTONOMY TO THE NOTION OF GOD

Is there any place for God in this picture of human autonomy, and if so, what would it be? A first answer assumed by many is: There is none, God is left outside. It is like Plato vs. Protagoras: No longer God, but humans are the measure of all things, deciding independently what is right and wrong.

A second, theological answer has been given by Aquinas and others.[6] As the creator of persons which are a source of value and valuable themselves, God is also the creator of all values entirely. He constitutes freedom in a deeper way, enabling us to make moral decisions and do what we recognize as morally good. Even more, he also provided a fundamental moral criteria or moral insight to be found in every human's soul, religious or not. This answer might be compatible with

[5] It is possible to extend the scope of obligation to all living beings. Cf. Korsgaard 2004.

[6] See the contribution of C. Schröer in this volume for an elaborated account of this approach.

a Kantian account (in this case the criteria is regarded the CI), however, it is not what Kant would say.

Once we turn to Kantian ethics, the *Religion within* suggests a third and different answer. There is a very surprising remark in the Preface, where Kant writes:

> This idea [of a moral, most holy and omnipotent being] is not ... an empty one; for it meets our natural need, which would otherwise be a hindrance to moral resolve, to think for all our doings and nondoings taken as a whole some sort of ultimate end which reason can justify. What is most important here, however, is that this idea rises out of morality and is not its foundation; that it is an end which to make one's own already presupposes ethical principles. (Rel. 6:5)

What exactly is 'a hindrance to moral resolve'? This is perhaps the most decisive part in Kant's ethical works, where it is not entirely clear if he is not close to introducing another motivational factor apart from the moral law on a different level.

God is introduced here as the provider of the highest good, a dimension of hope. If the common world we strive towards is only wishful thinking, an empty fantasy which will never become real, this could be an obstacle to moral decisions. God guarantees that sooner or later everybody will reach happiness (*eudaimonia*), not unconditionally, but according to her or his character, while he guarantees as well that we finally become those virtuous persons we are striving at. Yet this is not a different motivational ground for morality, as Kant instantly adds. If there were no hope for heaven or eternity, we are still supposed to act morally and be able to. For the Categorical Imperative to make sense, it must be assured the Kingdom of Ends is realized, sooner or later, as a community of rational beings in which nobody is using other persons as a mere means to his or her ends any longer. Since humans seem incapable to ever achieve this goal fully, God is the only guarantor that our moral efforts are not in vain. However, this dimension of hope is not regarded as a motivating factor for adopting ethical principles in the first place.

As I have tried to argue, the several sources of motivation outlined initially are actually only one: Valuing myself as a free human being, living in a world of other beings just like me. Referring to the title of my contribution, I do not think moral conflicts are taking place between compulsion and attractiveness, at all. They all come down to the question which attraction to follow, which compulsion to bow to, and why.

BIBLIOGRAPHY

Hare, John. 2005. 'Review of Linda Zagzebski, *Divine Motivation Theory*', *Notre Dame Philosophical Reviews*, 2.13

Kant, Immanuel. 1997. *Groundwork of the Metaphysics of Morals*, ed. by Mary Gregor (Cambridge: Cambridge University Press) [G]

Kant, Immanuel. 1998. *Religion within the Boundaries of Mere Reason And Other Writings*, Allen Wood & George di Giovanni (eds) (Cambridge: Cambridge University Press) [Rel.]

Korsgaard, Christine. 1996. *Sources of Normativity* (Cambridge: Cambridge University Press)

Korsgaard, Christine. 2004. 'Fellow Creatures: Kantian Ethics and our Duties to Animals', *Tanner Lectures on Human Values* 24: 77-110

Schwartz, Maria. 2006. *Der Begriff der Maxime bei Kant. Eine Untersuchung des Maximenbegriffs in Kants praktischer Philosophie* (Berlin: LIT)

Zagzebski, Linda. 2004. *Divine Motivation Theory* (Cambridge: Cambridge University Press)