

A quantitative model for a theory of justice. Part II: the extended quasi-maximin principle and moral psychology

Stijn Bruers

Abstract

In this article, we will look for ethical principles or moral intuitions that can overrule the quasi-maximin theory (QMM) and cannot be derived from the Rawlsian original position. One intuition is the ethics of care based on intersubjective empathy. To avoid inequality and discrimination in this ethics, we need a kind of “tolerated choice equality”. A second moral intuition is the deontological (Kantian) ethics of the basic right: the right not to be treated as merely means to an end. By looking at moral dilemmas, we demonstrate that this basic right often overrules the QMM-principle. We will present a mathematical equation that unifies this basic right principle with the QMM-principle. The advantages of this “extended QMM-theory” will be discussed. Finally, we postulate that the symmetry principle behind the QMM-formulation, the tolerated choice equality and the basic right equality are the three most important principles in a theory of equality.

Introduction

In the previous article we encountered the quasi-maximin principle (QMM) that is derived from the Rawlsian veil of ignorance (Rawls, 1971) with a high but not maximal risk aversion, and from the equality principle with a low but not zero need for efficiency. The mathematical formulation is described by the following. Suppose there are two situations X and Y. Situation X contains N individuals with qualities of life ordered as $x_1 \leq x_2 \leq \dots \leq x_N$. Person 1 is the worst-off, person N has the highest quality of life. Situation X can be described by the ordered set $X = (x_1; x_2; \dots x_N)$. Similarly, situation Y consists of M individuals, with qualities of life $Y = (y_1; \dots y_M)$ and ordered as $y_1 \leq \dots \leq y_M$. The QMM-theory now states that situation X is preferred over Y if and only if the moral weight $W(X) \geq W(Y)$, or:

$$W(X) = \frac{\sum_{i=1}^N Q^{i-1} x_i}{\sum_{i=1}^N Q^{i-1}} \geq \frac{\sum_{j=1}^M Q^{j-1} y_j}{\sum_{j=1}^M Q^{j-1}} = W(Y).$$

The parameter Q is a small but positive constant. This is the central and most general mathematical inequality to describe the quasi-maximin theory.

This QMM-theory is a consequentialist theory, based on rational (cognitive) thinking. Now, as is well known in studies of moral psychology (Greene et al., 2001; Haidt, 2001; Cushman et al., 2006), the moral brain does not only contain cognition, but also a large portion of emotional and intuitive judging. This idea goes back to David Hume’s emotivism (Hume, 1978). Greene et al. (2004; 2008) propose a dual-process model in which intuitive emotional responses and controlled cognitive responses in moral judgments both play a sometimes competing role. Emotional responses are related to a Kantian deontological ethics based on duties and obligations (Kant, 1959), controlled cognitive responses are related to the consequentialist ethics (Mill’s utilitarianism (Mill, 1998) or the QMM-theory). As is shown in a lot of moral dilemmas (some of them will be encountered later on in this article), there is a general difficulty to keep rational judgments in line with emotional/intuitive ones.

Cushman et al. (2006) gave three examples of moral intuitions that often play a role in our judgments:

- The action principle: doing harm is worse than allowing harm, or harm by action is worse than by omission (see also the discussion on doing versus allowing harm in Howard-Snyder, 2007).
- The intention principle: harm inflicted by using someone as a means to a goal is worse than equivalent ‘collateral damage’ (harm as a foreseen side effect).
- The contact principle: harm caused by using close or physical contact is worse than equivalent but ‘impersonal’ harm (without close contact).

The latter contact principle is similar to Greene’s distinction between personal vs. impersonal moral dilemmas. In personal dilemmas, the harm occurs in an ‘up close and personal’ manner (Greene et al., 2001).

There is another moral intuition, that intentional harm (doing something on purpose) is worse than unintentional harm (doing something by accident). However, this difference is not relevant for our QMM-theory, as our theory only involves situations where we have to choose (intentionally) between options X and Y. When we choose situation Y, even though $W(X) \geq W(Y)$, then this choice is always an intentional harm (there simply are no unintentional choices.)

We note that quasi-maximin, although a rational theory, already incorporates in some subtle ways some moral intuitions. In particular the level of risk aversion and the need for efficiency cannot be derived from purely rational reasoning. These non trivial (not zero or one) values for risk aversion and need for efficiency result into a prioritarian theory that already meets some intuitions. In this sense, utilitarianism and maximin are even further away from our moral intuitions, as they give extreme values for risk aversion and efficiency. However, there are a lot of other intuitive judgments that are in conflict with and cannot be derived from QMM (nor from the Rawlsian veil of ignorance in general).

In this article we will look for moral intuitions and emotions that might overrule the QMM-theory. We will make a distinction between ‘strong’ and ‘weak’ overruling. The QMM-theory will serve as a basic underlying structure, a backbone. The following scheme gives a summary.

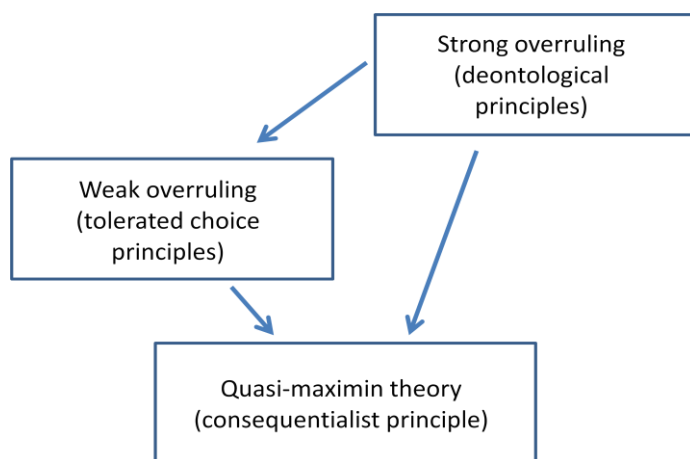


Figure 1: weak and strong overruling of the QMM-theory

Weak overruling of QMM: tolerated choice principles

Let us first look at the moral intuitions that weakly overrule the QMM-theory. In particular, we will look at the importance of intersubjective empathy and personal involvement. One can think here of the (feminist) ethics of care (Noddings, 2002) As a starting point, take the burning house dilemma. In a burning house, your child and an unknown child are threatened. Which child are you going to rescue first? Most people will say their own child. So your own child has a higher survival level.

In general, we often have difficulties being impartial, because we have strong emotions towards our relatives and friends. This partiality might conflict with consequentialist theories like QMM-theory. Also the abovementioned contact principle has the similar risk to influence our moral judgments in conflict with e.g. QMM-theory. We have stronger empathy with people that we know well, and stronger personal involvement when we have closer contact

with a person. This is typical in the ethics of care, which gives priority to personal relationships.

The question is: is this discrimination? Are not all people (and hence all children in the burning house) equal? Should we not toss a coin in order to fairly decide which child we are going to save? A specific formulation of the antidiscrimination principle can be derived from the QMM-theory. So if our choice in the burning house dilemma violates the antidiscrimination principle, it can also violate the QMM-principle.

Let's define discrimination as causing a disadvantage of an individual (or a group), based on a value-laden distinction between individuals (or groups), where the distinction refers to properties of the individuals (or groups) that are not deemed morally relevant in that situation. Or in other words: a person A discriminates B against C, if A believes that B has lower value than C (meaning C should have more rights, advantages or opportunities than B), where this value difference is derived by properties of B and C that are not morally relevant or are not an acceptable motive for the decisions and behavior of A.

The question is: what are morally (ir)relevant properties or criteria? If we put this in the QMM-framework, the answer is simple: morally relevant properties are all properties that are related to improving the quality of life of all individuals in line with the QMM-principle. For example, in the previous article we discussed desert based and resource based principles. So morally relevant properties are amongst other things: desert (contribution to the quality of life of others), effort, incurred costs and personal responsibility of the concerned people.

Now, the tricky point is that your child is not more deserving or responsible, simply because it is your child. Actually, the fact that it is your child is not important for QMM-theory. Of course, if you lose your child, your quality of life will be affected. But don't forget that the parent of the other child in the burning house will also feel sad when his child dies. The death of your child is as bad as the death of the other child, if we look at QMM-theory. The moral weights in both situations are the same ($W(X) = W(Y)$).

So do we have discrimination? In some sense yes, as we have just argued. There is an emotional inequality in our behavior towards different children. Yet, there can still be some subtle form of equality present. It is related to the words "value-laden" in the definition. What do we mean by that? Suppose the parent of the other child passes the burning house, and saves his child. There are two ways for you to react. You can say that what that person did was immoral, because your child has more intrinsic value and should have more rights. Or you can say that, although you regret that your child died, you accept and tolerate the choice of that person to save his child. This is what we can call "tolerated choice equality". You and the other parent are in some sense equal, and therefore your children inherit a tolerated choice equality, although there is an emotional inequality from your point of view.

Tolerated choice equality should be distinguished from a moral value-laden inequality, which results from a discriminatory ideology such as racism, sexism or speciecism (Ryder, 1975). These kind of ideologies are hierarchical dualisms (Plumwood, 1993), which can be characterized by one or more of the following characteristics (this is a little extension of the theory of Plumwood).

- 1) There is a deep gap and a radical exclusion between the two poles (the upper side of oppressors and the lower side of oppressed). Any overlapping between the two poles is denied.
- 2) The lower side is negatively defined: the oppressed lack the properties which are used by the oppressors to justify the oppression.
- 3) The lower side is marginalized: the oppressors do not show care and empathy. The personalities and needs of the oppressed are denied or scorned.
- 4) The lower side is homogenized, individual differences between people from the lower side are denied, by use of e.g. stereotyping,...

5) The lower side is unjustly criminalized, they are the scapegoats. The innocence of the lower side is denied.

6) The lower side is instrumentalized (objectified), they are used as tools, as means to the ends of the upper side. The intrinsic value of the lower side is denied.

When one or more of these criteria are satisfied, there is a moral value-laden difference, and we can speak of immoral discrimination. When there is tolerated choice equality, there cannot be an hierarchical dualism and hence no immoral discrimination.

In general, we have “tolerated choice principles” that weakly overrule the QMM-theory. It might be that a person has emotional difficulties in reaching QMM, but at least he should tolerate the choice of someone else if that is in line with QMM. So we should always at least accept the choice of someone who prefers X over Y if $W(X) > W(Y)$.

However, there are some stronger moral intuitions that cannot be put into tolerated choice principles. Some moral judgments that violate QMM will not be tolerated. The next section gives an example of such a strong overruling of QMM by a moral intuition.

Strong overruling of QMM: deontological principles

The famous trolley dilemmas (Thomson, 1985) are an interesting tool to demonstrate the importance of deontological judgments. In the first dilemma, referred to as the switch dilemma, a trolley is going to kill five people lying on the tracks. However, you can hit a switch so that the trolley takes a side track, to save the five people. Unfortunately, on this side track there is one person lying. In short: doing nothing results into the death of five people, pulling the switch results in the death of one person. Most people say that one should hit the switch (Petrinovitch et al., 1993), in line with the consequentialist ethics: both maximin and utilitarianism state that one person dead is better than five persons dead.

In the second dilemma, referred to as the bridge dilemma, again a trolley is about to hit five people. However, you are standing on a bridge, next to a heavy person with a heavy backpack. You can push the person from the bridge in front of the train in order to block the train. The heavy person will die, but the five persons on the track will be saved. In this case, most people would say that we are not allowed to push the person from the bridge. This is a deontological answer. By looking at brain activity, Greene et al. (2001) discovered that persons reacted more emotionally in this second dilemma, compared to the first.

In a third dilemma, called the hospital dilemma, five patients in the hospital need new organs in order to survive. However, no organs are available anymore. Is it allowed to kill a visitor and use his kidneys, liver, heart and spleen for transplantation to save the five persons? Also here, most people are deontologists: they are very reluctant to act, even if they save more lives.

What are the distinctions between those dilemma's, that can explain the different judgments? There are two possible explanations. As a first distinction, we need to ask whether the presence of the victim we're about to harm is necessary. Can we still flip the switch if the person on the side track was absent? Yes, our plan would still work. Can we save the five persons in the bridge dilemma if the heavy person on the bridge was absent? No, there is no one to push. Can we save the five patients in the hospital if no visitor was present? Clearly no. So, our first hypothesis is that we should not act and kill a person if the presence of that person is required in order to save the lives of others. But there is a second explanation, as mentioned by Waldmann and Dieterich (2007). They look at the point within the causal model that is targeted by the intervention. There are three possibilities, as explained by the three dilemmas. In the first dilemma, we influence the path of the “agent” (the trolley): we redirect the trolley and send it to the one person on the side track. The locus of our intervention is at the trolley, and the trolley will eventually influence (the path of) the victim. In the second dilemma, we directly influence the path of the “patient” (the victim, i.e. the heavy person on

the bridge): we send the victim to the agent by pushing him. The locus of our intervention is at the victim. So in the first dilemma, we send the agent to the victim, and in the second we send the victim to the agent. In the third dilemma of the hospital, the action immediately influences the victim (the visitor) and kills him. So also in this dilemma we influence the path of the victim, but even more directly than in the bridge dilemma: the person who kills the visitor is himself the agent.

So the second hypothesis is that we should not act and kill a person if our action influences the path of the victim (if the locus of our intervention in the causal chain lies at the victim).

We now have two hypothesis. They are not equivalent, as the following dilemma, called the loop dilemma (Waldmann et al., 2007), demonstrates. As in the switch dilemma, you are standing on a switch, five people are on the main track, and one (heavy) person is on the side track. The difference is that the side track is small and comes back to the main track at a point before the five persons. So the tracks make a small loop. If you don't hit the switch, five people will get killed. If you turn the switch, the heavy man on the side track will get killed by blocking the train. However, if the heavy person was not on the side track, the trolley will move one and enter the main track again to kill the five persons. The presence of the person on the side track is therefore necessary in order to save the five people. And turning the switch does not directly influence the path of the victim: the train is sent to the victim, not vice versa. The condition of hypothesis 1 is met, the condition of hypothesis 2 is not.

According to Hauser et al. (2008), people are divided on the acceptability of turning the switch in this loop dilemma (roughly half of the respondents stated that one should turn the switch). However, in another empirical study, Waldmann et al. (2007) found no significantly different judgments between the loop and the switch dilemma. Yet, in the study of Hauser et al., there was still a significant different response between the loop dilemma and the "loop plus rock" dilemma. In the latter dilemma, the person on the side track of the loop is standing in front of a heavy rock. So this time, the rock will block the train, and the person is "collateral damage". I.e. the presence of the person on the side track is not necessary anymore to save the five people. In the "loop plus rock"-dilemma, more than 70% of the people judged that one should turn the switch. This is significantly more than in the loop dilemma.

So who is right? Our first hypothesis seems to be correct to some level: the required presence of the victim does seem to play a role, even in the loop dilemma (as can be seen by the difference with the "loop plus rock" dilemma). But also the second hypothesis (the locus of intervention) is valid.

No matter what the validity of the two hypotheses is, we will now simply reformulate these hypotheses in one deontological (Kantian) principle: a subject has the basic right not to be used as merely means to someone else's ends. In the switch dilemma, the person on the side track is "collateral damage" and not a means to save the five people (Thomson, 1985; Waldmann et al., 2007; Moore et al., 2008). In the bridge and hospital dilemmas however, the victim is being used as a trolleystopper and a set of organs respectively.

The formulation of the deontological basic right is rather vague:

-What do we mean by use as "merely" means? Are all persons who are being used as means without their informed voluntary consent, used as "merely" means, or are there other criteria?

-What do we mean with "ends"? Are vital needs at the same level as luxury ends?

-What is the exact correspondence between the basic right and the two hypotheses formulated above?

We will not go into detail about these questions. What we want to stress here, is that there exists at least one principle, a deontological (Kantian) right, that violates the consequentialist (QMM) principle and cannot be derived from e.g. the Rawlsian veil of ignorance. Furthermore, this deontological principle is very strong: we do not tolerate the choice of the

people who violate this principle. That is why we speak of a “basic right”, a principle that strongly overrules the QMM-principle.

Can we incorporate this new principle in our mathematical formulation? We suggest the following possibility: apart from their qualities of life, all individuals have a “basic right parameter” R_i^X , which is zero if the basic right of person i in situation X is not violated, and very large if her basic right is violated. The quantity R_i^X can also take different values, depending on the answers of the above questions: how “merely” is the use as means, what are the “ends”, where is the locus of intervention,...? With these basic right parameters, we can construct a new moral weight

$$W_R(X) = - \sum_i^N R_i^X.$$

When the basic right of person i is violated, the value R_i^X should be very large, but not infinite: there might be a threshold value, above which one prefers the consequentialist outcome. For example: when there are say a billion people on the main track (all threatened), one might be tempted to push the heavy man from the bridge in order to save those billion people. Another reason why R_i^X should not be infinite, is because it is difficult to count with infinities, as infinity plus infinity equals infinity. So when we have to choose between situation X where one basic right is violated and situation Y where the basic rights of two persons are violated, we should be able to conclude that the situation X is preferred over Y.

Having said this, we can now present the new mathematical formulation of the QMM-principle extended with the deontological rule of the basic right. Situation X contains N individuals with qualities of life ordered as $x_1 \leq x_2 \leq \dots \leq x_N$. Situation Y consists of M individuals, with qualities of life $y_1 \leq \dots \leq y_M$. Then situation X is preferred over Y if and only if:

$$\frac{\sum_{i=1}^N Q^{i-1} x_i}{\sum_{i=1}^N Q^{i-1}} - \sum_i^N R_i^X \geq \frac{\sum_{j=1}^M Q^{j-1} y_j}{\sum_{j=1}^M Q^{j-1}} - \sum_j^M R_j^Y.$$

The “extended” moral weight of option X thus reads:

$$W_{ext}(X) = W_{QMM}(X) + W_R(X) = \frac{\sum_{i=1}^N Q^{i-1} x_i}{\sum_{i=1}^N Q^{i-1}} - \sum_i^N R_i^X.$$

This is the most complete equation of the consequentialist QMM-principle combined with the deontological basic right principle. With this formulation, combined with the weaker “tolerated choice” principles, we can get a very rich ethical system that is compatible with a lot of moral intuitions that a lot of people have. Let us call this the “extended quasi-maximin theory”. As the right principle is very strong (the terms R_i^X can be very large), it cannot be overruled by the weaker tolerated choice principles (like e.g. the ethics of care). In other words, if I were in a burning house dilemma and had to choose between my child and an unknown child, I would save my child. But that doesn’t mean that I’m allowed to use an unknown child as organ donor to save my child from kidney damage. So the strong deontological principle overrules the weak principles and the QMM-principle see Figure 1).

There are of course other principles or intuitions that are not included in this theory (and that the author does not share). There are for example principles based on metaphysical religious beliefs about souls,... (think about the problem of euthanasia and abortion: according to our extended QMM-theory, these practices are not always immoral). Also some ideologies do not fit with our theory. In particular, our theory, applied to all sentient beings (as should be the case: all sentient beings can experience a quality of life) lead to e.g. rather radical shifts in diets (to some kind of veganism). Most people don’t have the idea that animals have rights. These people have a speciest (Ryder, 1975) and carnist ideology (Joy, 2001), with an ingroup-outgroup distinction that is often perceived by them as natural or instinctive.

However, these ideologies are not as instinctive as are our instinctive moral judgments mentioned above. The ideologies are more culturally determined, and can change more rapidly than some of our true moral instincts. As in racism and sexism, the dividing line between ingroup and outgroup can be influenced and is vulnerable to change. The fact that a basic right can overrule QMM, is related to a strong intuition. But the questions who gets the basic right and who should be counted in consequentialist calculations, can be influenced by ideologies and cultures. These are not as “hard wired” in our brains.

The action principle

Looking at the three intuitive principles studied in Cushman et al. (2006), we see that the contact principle is incorporated in the “tolerated choice”-principles that weakly overrule the QMM-principle, and the intention principle is nothing but the deontological principle of the basic right, that strongly overrules the QMM-principle. But what about the action principle? This has to do with the difference between doing versus allowing harm. This difference is not always clear: it is difficult to pinpoint exactly what makes the difference between a doing and an allowing (see Howard-Snyder, 2007 for an elaborate discussion).

Let us study a famous example: the difference between pushing someone into the water in order to kill him, versus letting drown someone who fell into the water. The first is a doing, the second is an allowing. In our mathematical formulation, we have four situations. In the first situation, everyone is on shore and nothing happens. By symmetry we then have to assume (all else equal) that both persons have a similar well-being represented by the couple $X = (1; 1)$. In the second situation, one person is pushed into the water and drowns. This person then gets a well-being equal to zero, whereas the other person gets a little increase in well-being (he feels a little happier, otherwise he would not push someone into the water to kill him). So we get $Y = (0; 1 + a)$, with a a positive number. In the third situation, which is the “allowing”, the first person is drowning and the other person is standing on shore doing nothing to help him. So the well-beings can be represented as $Y' = (0; 1)$. In the fourth situation, the person on shore jumps into the water in order to save the drowning person. As this first person is always a bit reluctant to jump into the water (he would not jump if no-one was drowning), his well-being will decrease a little, and the well-being of the saved person will increase (but will still be lower than one, as he did not prefer to fall into the water). So this situation is represented as $X' = (b; 1 - c)$, with $0 < b < 1$ and $0 < c < 1$.

We know that $W(X) > W(Y)$, and $W(X') > W(Y')$. Now we can define a quantity

$$\Delta = W(X) - W(Y) - (W(X') - W(Y')).$$

If $\Delta > 0$, than that actually means that the doing is worse than the allowing (i.e. $W(X) - W(Y) > W(X') - W(Y')$, so the drop in moral weight in the doing is higher than in the allowing). Δ is positive if $1 - b > Q(a - c)$ when $b < 1 - c$ or $c > Q(a + b - 1)$ when $b > 1 - c$. We see that if Q is very small (as is the case in QMM-theory), Δ is almost always positive. In other words, the fact that throwing someone in the water is worse than letting someone drown, can already (at least in some circumstances) be derived from the QMM-principle.

Whether the other examples of doing vs. allowing can be derived from the QMM-principle itself or constitute a weak overruling (with a tolerated choice) is not clear yet and needs to be further studied. The author is not aware of the existence of a doing vs. allowing example that strongly overrules the QMM-principle. So or above expression of the extended moral weight should not yet be adapted to incorporate an intuition about doing versus allowing.

Three kinds of equality

In this final section, we summarize our extended QMM-theory, by postulating that there are three most fundamental principles in a theory of equality.

- 1) Impartiality towards individuals from behind the veil of ignorance. This impartiality leads to important symmetries in our mathematical formulation. First, if there are N individuals, we have equal probability to be each of those individuals. So the probability to be person i equals $1/N$. Second a situation where two persons have qualities of life respectively x_1 and x_2 , equals a situation where the qualities of life are reversed (person 1 has x_2 , person 2 has x_1). This property allowed us to write the situations as ordered sets $X = (x_1; x_2; \dots x_N)$ with $x_1 \leq x_2 \leq \dots \leq x_N$.
 - 2) Tolerated choice equality: even though we would save our own child in the burning house dilemma, we would tolerate the choice of someone else who has saved the other child.
 - 3) Basic right equality: $R_i^X = R_j^Y$ if the use as a means of subject i in situation X is similar to the use of subject j in situation Y (i.e. for the same kind of ends,...). We let the possibility open that a basic rights violation of one person can be worse than the violation of another person, depending on differences in e.g. mental capabilities between those persons.
- These three principles of equality are at the foundations of resp. QMM-theory and the weak and strong overrulings of QMM.

Who has to follow the extended QMM-theory?

The question now arises who has to obey (follow, implement,...) this new theory of justice? It is clear that society as a whole (and in particular political, economical and legal bodies) has to try to follow the extended QMM-theory, as political or legal bodies can be impartial and problems with tolerated choice principles can be avoided (there is no “other” outside of society, so it is meaningless to tolerate the choice of this non-existent “other”).

But what about individuals? When do they have to obey the extended QMM-theory? First, although this theory applies to all moral patients (including animals, babies,...), only moral agents have to obey this theory, as they are the only persons who can understand the principles. The question we want to ask is: Does a person still has to obey the extended QMM-theory, even if no-one else does so? We would argue that the answer is “yes”, in line with the universal golden rule or the categorical imperative, which can be formulated in this context as: Act according to those principles that, if all moral agents would act according to those same principles, the situation of extended QMM is most closely approached.

This categorical imperative puts both strong and weak constraints on a person’s behavior. The strong constraint can be seen in the famous prisoner’s dilemma (Rapoport, 1966). In this dilemma, there are three situations with two persons involved. In the first situation, the two persons cooperate, and their qualities of life will be $X = (3; 3)$. In the second situation, only the first person cooperates, which will give the qualities of life $Y = (1; 4)$. So the person who decided to cooperate, is worst off. In the third situation, no-one cooperates, and both will end up with qualities of life $Z = (2; 2)$. It is clear that $W(X) > W(Z) > W(Y)$. In other words, if everyone cooperates, QMM will be realized. According to the categorical imperative, we should therefore always cooperate, even if the other does not cooperate. Notice that if the other would not cooperate, we end up in the situation Y , which has the lowest moral weight. So following the categorical imperative is sometimes not easy. What individuals have to do is twofold: first they should decide for themselves to cooperate, and second they should put up political structures that enforces cooperation on all other moral agents. (For an iterated prisoner’s dilemma, one can use the evolutionary successful tit-for-tat strategy: the first move should be cooperation, and the following move is always the same as the opponents previous move. See Axelrod, 1984).

Following the categorical imperative, on the other hand, can also relief ourselves from very heavy duties. For example: I can try to erase extreme poverty on my own by donating nearly all of my money to charity. This might be the implication of the extended QMM-theory, if the other richer persons refuse to donate anything. However, if all the richer persons donated a

substantial (but not extremely high) portion of their incomes to charity, the global QMM situation might be approached as well. Therefore, following the categorical imperative, I should donate a very substantial amount of my income to charity, but not such an extremely high amount that would drive me close to extreme poverty myself.

References

- Axelrod R., 1984, *The Evolution of Cooperation*. Basic Books.
- Cushman F., Young L., Hauser M., 2006, *The Role of Conscious Reasoning and Intuition in Moral Judgment. Testing Three Principles of Harm*, *Psychological Science*, 17, nr.12:1082-1089.
- Greene J., Sommerville R., Nystrom L., Darley J., & Cohen J., 2001, *An fMRI investigation of emotional engagement in moral judgment*, *Science*, 293:2105–2108.
- Greene J., Nystrom L., Engell A., Darley J., Cohen J., 2004, *The neural bases of cognitive conflict and control in moral judgment*. *Neuron* 44:389-400.
- Greene J., Morelli S., Lowenberg K., Nystrom L., Cohen J., 2008, *Cognitive load selectively interferes with utilitarian moral judgment*, *Cognition* 107:1144-1154.
- Haidt, J., 2001, *The emotional dog and its rational tail: A social intuitionist approach to moral judgment*, *Psychological Review*, 108:814–834.
- Hauser M., Young L., Cushman F., 2008, *Reviving Rawls' Linguistic Analogy: Operative principles and the causal structure of moral actions*. In *Moral Psychology*, Vol 2, Cambridge: MIT Press.
- Howard-Snyder F., 2007, *Doing vs. Allowing Harm*, *Stanford Encyclopedia of Philosophy*.
- Hume D., 1739/1978, *A treatise of human nature*. 2d Edition, Oxford: Oxford University Press.
- Joy M., 2001, *From carnivore to carnist: Liberating the language of meat*, *Satya*, 8(2), 26-27.
- Kant I., 1785/1959, *Foundation of the metaphysics of morals*. Indianapolis: Bobbs-Merrill.
- Mill J.S., 1861/1998. *Utilitarianism*. New York: Oxford University Press.
- Moore A., Clark B., Kane M., 2008, *Who shalt not kill?: Individual differences in working memory capacity, executive control, and moral judgment*, *Psychological Science* 19:549-557.
- Noddings N., 2002, *Starting at Home: Caring and Social Policy*, Berkeley, CA.: University of California Press.
- Petrinovich L., O'Neill P., Jorgensen M., 1993, *An empirical study of moral intuitions: Toward an evolutionary ethics*, *Journal of Personality and Social Psychology* 64:467-478.
- Plumwood V., 1993, *Feminism and the Mastery of Nature*, Routledge, 1993.
- Rapoport A., 1966, *Two-person game theory: the essential ideas*. Ann Arbor, University of Michigan Press.
- Rawls J., 1971, *A Theory of Justice*, Cambridge: Harvard University Press, rev. ed. 1999.
- Ryder R., 1975, *Victims of Science: The Use of Animals in Research*. Davis-Poynter.
- Thomson J., 1985, *The Trolley Problem*, *Yale Law Journal* 94:1395-1415.
- Waldmann M. and Dieterich J., 2007, *Throwing a bomb on a person versus throwing a person on a bomb. Intervention myopia in moral intuitions*, *Psychological Science*, 18, nr.3:247-253.