

# *Universalism and Obligation*

*- Is Ernst Tugendhat's concept of an emotional obligation  
to universalism defensible?*

*Andreas Eriksen*

Thesis submitted for the Master of Arts Degree

Supervisor: Professor Christel Fricke

Department of Philosophy, Classics, History of Arts and Ideas

May 2009



**CULCOM**

Cultural complexity  
in the new Norway

## *Acknowledgements*

Christel Fricke has been my supervisor and I must thank her for many helpful discussions. She has made me think more rigorously about the problem and offered much advice that was vital to my thesis.

I would also like to thank Harald Grimen, Eivind Kirkeby and Anders Molander for having commented upon my work at different stages. This has improved the final result in many ways. I should also extend thanks to Ketil Bonaunet, whose lectures on Kant in Tromsø inspired me to write this thesis.

I am grateful for the scholarship that I have received from CULCOM. Affiliation with this interdisciplinary program has provided me with a greater range of viewpoints on the matter through seminars and project discussions. Thanks are due to my CULCOM-supervisor Odin Lysaker for encouraging and instructive remarks.

Much appreciation goes to my parents, who have offered both invaluable academic guidance and much personal support.

And lastly, I cannot thank my soon-to-be wife Agnete enough. She has she taken great care of our newborn Ludvig and provided me with the necessary time to work. Her love, support and understanding made this project possible.

# Contents

<i>Introduction</i> .....	3
<i>Terminology</i> .....	6
<b>Part 1</b> .....	<b>7</b>
1. <i>Stating the Problem</i> .....	7
1.1 Method.....	7
1.2 Modernity .....	10
2. <i>What is Morality?</i> .....	16
2.1 Moral Obligation .....	17
2.2 Moral Goodness .....	21
2.3 Emotions of Self-Assessment and Obligation .....	25
3. <i>Why Universalism?</i> .....	28
3.1 Pioneering Investigation .....	29
3.2 Agreement as Pre-Condition for Collective Actions .....	31
3.3 Agreement or Force?.....	33
3.4 A Reciprocal Demand.....	37
4. <i>Justice and Discrimination</i> .....	40
4.1 Cooperation and Universalism .....	40
4.2 Equal Respect and Unequal Distribution .....	43
5.1 Why Justify?.....	47
5.2 Why Accept a Norm?.....	49
5.3 Justification of a Universalist Concept of Morality .....	52
5.4 Summary of Moral Justification .....	58
<b>Part 2</b> .....	<b>61</b>
6. <i>Sanctions and Obligation</i> .....	62
6.1 Dissolution of Obligation?.....	62
6.2 Theorist and Agent.....	67
6.3 Conceptual Priority, Not Sequential.....	71
7. <i>Contractarianism vs. Wishful Thinking?</i> .....	74
7. 1 Universalism as a "Subjective Ideal" .....	74
7.2 Who wants Justification? .....	75
7.3 Impoverished Topography .....	79
7.4 Emotional Substitutes?.....	81
8. <i>Conclusion</i> .....	85
<i>Literature</i> .....	89

## *Introduction*

This essay is a discussion of Ernst Tugendhat's moral philosophy. Over the last thirty years or so he has been investigating how we can justify moral norms without presupposing that agents believe in something "higher". That is: Can moral norms be understood independent of religion, tradition and other authorities that take the decision on what is right and wrong away from the agents themselves? Is it possible to justify moral norms by recourse to ordinary interests?

Tugendhat's starting point is the idea of moral obligation. Moral norms tell us what we are supposed to do, no matter what our preferences happen to be. How is this kind of normative necessity to be understood? According to Tugendhat we have to follow the norms in order to avoid a special kind of sanctions, the "inner" sanctions of conscience. We experience guilt insofar as we have transgressed a norm that we think is justified. So the concepts of obligation and justification are interdependent: We are only obligated to norms that we consider justified, and the justification of moral norms aims at showing why we have a reason to place ourselves under an obligation to be moral.

In what way do the norms have to be justified? The norms cannot merely be to the advantage of each agent, according to Tugendhat, they must satisfy a standard of moral goodness as well. He claims that moral obligation is dependent on norms that are *equally* justified to all affected by them, and argues that the standard of universalism as explained by Kant is the only defensible foundation for our norms: "*Handle so, daß du die Menschheit, sowohl in deiner Person als in der Person eines jeden anderen, jederzeit zugleich als Zweck, niemals bloß als Mittel brauchst*" (Kant 1785:429). Tugendhat summarises this second formulation of the categorical imperative in a statement that is instantly comprehensible: "Instrumentalisiere niemanden!" (1993, p. 80) A norm that does not satisfy this criterion is unjustified, then.

The first part of this thesis will present and discuss the way Tugendhat has arrived at these conclusions. I am aiming at understanding his project and to discuss how difficulties that occur can be resolved. As is mentioned in the introduction to the critical reader *Ernst Tugendhat: Einwände und Erwiderungen*, the prolonged debate that he has had with himself has caused doubts as to whether one can speak of *the* moral philosophy of Tugendhat (cf. Scarano and Suárez 2006, p. 8). In one of his very

last writings on morality Tugendhat saw it fit to write this characteristic yet frustrating little footnote: "Hingegen halte ich vieles, was ich früher über moralische Begründung gesagt habe, inzwischen für falsch" (Tugendhat 2003, p. 73n1). But this is in my opinion symptomatic of his tendency to exaggerate differences that exist between the various accounts of moral justification that he has given. It will be demonstrated that he has a unified moral theory in spite of all the revisions.

Tugendhat's moral philosophy has not received much attention outside the German language community. According to Tugendhat himself, his central "Drei Vorlesungen" is known only indirectly and distortedly to English readers through a denunciation written by Habermas (cf. Tugendhat 1997a:8). It is fair to say that none of his subsequent books have fared any better internationally. Hopefully, the first part of this thesis will convey that this is unfair, given that there is rich and interesting scope of ideas in his writings that are relevant to current debates on moral issues on for example the nature of normativity and justification of norms.

The second part of the thesis is a discussion of criticism that has been launched against Tugendhat from two opposite camps. Jürgen Habermas is the most prominent advocate of discourse ethics, a theory of moral justification that favours the same criterion as Tugendhat for deciding when a norm is justified. Like Tugendhat, Habermas claims that a norm has to be equally good for all, but he has arrived at this criterion by a different path. He thinks that a concept of "communicative reason" is necessary in order to justify this criterion. The norms derive their "*Sollgeltung*" by satisfying the standards of this concept of rationality, meaning that they are acceptable in a free discourse. Therefore, he does not think that emotional sanctions can help us understand the concept of moral obligation. Sanctions are something *secondary* to the obligation that we have as members of a rational discourse. The idea is that we are bound by the "zwanglosen Zwangs guter Gründe" (Habermas 1991, p. 151). Habermas would say that Tugendhat has a justifiable normative standard, then, but at the same time claim that obligation to the norms that satisfy this standard cannot be explained on Tugendhat's premises.

The contractarian theory that has recently been developed by Peter Stemmer in *Handeln zugunsten anderer* (2000) faces Tugendhat with an opposite critique. Stemmer agrees with Tugendhat that moral obligation must be understood in light of emotional sanctions, but he does not think that norms have to satisfy some Kantian

standard of moral goodness: "Die Idee einer 'säkularisierten' Moral, die sich inhaltlich mit der herkömmlichen theonomen Moral deckt, war wohl von Anbeginn an Wunschdenken, ein schöner Traum der Aufklärung" (2000, p. 286). Instead, the norms have to be *beneficial* for each agent in such a way that one could imagine that they were the result of a contract made out of self-interest. We can have no obligation to moral respect towards those that cannot hurt our interests, so moral respect is internal to a limited moral community of agents of equal strength. According to Stemmer, this is all we can justify philosophically when all transcendent moral authorities are gone.

Tugendhat's theory is countered with two difficulties, then: First, he wants to understand moral legitimacy in terms of universalism, but is accused by Habermas of having an inadequate explanation as to why we are obligated to this ideal. Second, he wants to justify our moral norms by recourse to our ordinary interests, but according to Stemmer this can never result in an obligation to equal moral respect towards all. It is this tension that is to be explored in the second part of this thesis. Is it possible to understand our obligation to the principle of universalism in terms of reactive attitudes and interests?

### *Terminology*

A brief overview of some terminological issues is in order before we begin, which should hopefully be done without pre-deciding any substantial issues. I will use the term 'interests' simply to denote the content of the will of the agents. Not some "authentically free will", just the will that is signalled in saying yes or no to an option. More will be said on this matter, but for now it is enough to recognise that interests function as a placeholder where there were previously unified beliefs in some "higher" moral authority, like God or a legend.

I will speak of Stemmer and David Gauthier's 'contractarianism' instead of 'contractualism'. This is done in order to separate their Hobbesian position from the contractualist projects that stand in the tradition from Locke, Kant and Rousseau (e.g. John Rawls and T.M. Scanlon). The non-Hobbesian way of discussing morality in terms of contracts has a principle of reciprocity built into the setting of the hypothetical contract. Contractarianism, on the other hand, is a way of seeing norms as legitimatised by being in our self-interest; they do not have to satisfy a criterion of equal respect for all affected by the norm.

Lastly, 'universalism' is used in the special sense of referring to the content of Kant's categorical imperative: All must be treated with equal respect, regardless of natural or cultural differences like race or religion. 'Universalism' is used instead of Tugendhat's varying ways of naming the standard. He usually speaks of "symmetry" or "equality", and these are egalitarian notions that have until recently been treated separately from an explanation of why we should respect the will of *all*, not just a limited society. But his latest understanding of morality interprets the basic principle as a criterion that involves both *equality* within the community and that the community should include *all* possible cooperative beings (cf. Tugendhat 2006b, p. 309). I could of course call this standard "symmetry", except that Tugendhat has indicated that this should be understood as a descriptive notion and not a prescriptive standard (cf. 2007b, p. 147). But occasionally referring to his position by his own term "symmetric contractarianism" seems appropriate nevertheless.

## *Part I*

### *1. Stating the Problem*

In this chapter I am seeking to find out how Tugendhat himself understands the problem of justification of our moral practice. This must be answered with regard to two separate issues. The first issue is how Tugendhat envisions the role of the philosopher. How can the philosopher contribute to questions on morality? This is a question of *method*. I will take a short look at how his methodological stance contrasts with the way his critics view the task of moral philosophy, since this will bring the problem of this thesis further into the light.

The second issue is how Tugendhat himself understands the society's need for a philosophical account of moral justification. That is, why do we need a theory that tells us how we can justify moral principles by recourse to ordinary interests? This is the problem of *modernity* - as Tugendhat uses this concept. He has been unduly vague on this issue, so perhaps it can be recast in more definite terms.

#### *1.1 Method*

Clearly, the role one envisions for the philosopher is going to affect what kind of results one can expect. The issue of philosophical methodology has been of great concern to Tugendhat; he has forcefully argued that the traditional ontological philosophical questions must be recast as analytical reflections upon fundamental concepts and their anthropological basis (see esp. Tugendhat 1976 and 1989a). I will not concern myself with the details of this meta-philosophy, but only try to get a general picture. His account of morality must be viewed as an application of his more comprehensive theory of philosophy itself.

According to Tugendhat, a philosophical account will go astray if it tries to do more than conceptual clarification: "Die Philosophie kann nicht mehr tun, als ein vorhandenes Vorverständnis in seinen Voraussetzungen adäquat zu analysieren; sie hat keinen eigenen, extramundanen Bezugspunkt" (Tugendhat 1993, p. 28). We cannot expect philosophy to arrive at some truth beyond what is already implicit in human practice and understanding. What the philosopher can do is reflect upon the expressions and the attitudes of the practice that he himself is a part of, describing how it is possible that we think and act in the way we do. He can explicate the



meaning and interrelations of concepts, but he cannot construct a new meaning that is independent from the pre-reflective understanding that we already have of things.

Tugendhat therefore claims that the moral practice must be *understood* and not altered. Philosophy is there to capture the concepts that are used by agents in a more systematic way: "Für mich besteht die Aufgabe der philosophischen Reflexion auf Moral, darin, das Phänomen aufzuklären, so wie es ist, und nicht darin, sich zu fragen, wie es nach irgendwelchen - 'kognitivistischen' - Maßstäben aussehen müßte" (Tugendhat 2006b, p. 297). Although clearly reminiscent of Hegel's philosophical program as stated in the introduction to the *Rechtslehre* - where philosophy is said to be „ihre Zeit in Gedanken erfaßt“ (1821, p. 26) - the similarities with Hegel when it comes to doing moral philosophy end there. To Hegel it is important that the individual is at one with society and its moral norms; if individuals experience a need for an explicit justification this is said to be symptomatic of anomic and destructive sides of society.

Tugendhat, on the other hand, sees the individual's quest for moral justification as a healthy and important endeavour. The philosopher can pave the way for this undertaking, tidying up confusions and showing what our practical decisions actually entail. We have to understand the critique from both Habermas and Stemmer in light of this aim of Tugendhat's. The issue is whether or not the practical options that we face have been analysed in an adequate manner. As we shall see in Part 2, both of the critics think that he has given moral agents one option too many. Habermas thinks giving agents the option of choosing whether to be moral or not destroys the possibility of obligation. While Stemmer, on the other hand, thinks that making morality into an option of behaving in accordance with universalism involves a concession to irrational ideals. Can these misgivings about Tugendhat's theory be traced back to methodological presupposition?

Yes, since both critics want to elucidate how we should perceive moral norms according to their respective standards of rationality. Both authors aim at "Klärung der rationalen Grundlagen von ... Handeln" (Habermas 2004b, p. 355), that is, to discover "worin die unterstellte Vernünftigkeit des moralischen Handelns liegt" (Stemmer 2000, p. 10). The difference between the two critics lies in their respective concepts of reason: Habermas has a *communicative* concept of reason, meaning that norms are rational only to the extent that they could be the objects of a consensus in

an "ideal discourse situation". This basically implies that rationality is defined in terms of a discourse where only epistemic reasons rather than brute force count. Stemmer's account goes in the opposite direction, since he is saying that norms are rational relative to strength and interests. What all this implies will become clear when discussing their accusations, but we can already perceive differences to Tugendhat. Both critics start off with an explication of a standard of rationality in light of which they understand morality: Habermas sees moral norms as validated through his consensus theory of validity, while Stemmer has egoistic self-interest as the sole criterion of legitimacy.

Knowing Tugendhat's actual aim is central not only in order to understand his project, but also in order to evaluate the criticism he receives. It makes clear why it the primary interest should not be to investigate whether or not his theory fails to live up to the standards of rationality that belong to the alternative accounts. When Habermas claims that Tugendhat's account "verrät einen kognitiven ... Sinn der Sollgeltung moralischer Normen", for instance, we know that this is lethal only once it has been demonstrated that obligation to moral norms must be understood only in terms of rationality. But as the quote above reveals, Tugendhat thinks a philosophical account of moral obligation should not let standards of rationality predetermine the outcome. The issue of the second part of this thesis is to decide whether this is a defensible position.

So Tugendhat's aim is to present an understanding of the nature of our moral norms and their justification, and our interest in Habermas and Stemmer is due to their challenge that Tugendhat has an incoherent understanding of moral obligation. The criticism is interesting because of the partial overlap Tugendhat's theory has with the competing accounts. What I have tried to indicate is that the basic differences can be explained by tracing them back to the divergence of their aims. To Habermas and Stemmer it is more important to illuminate how norms could be legitimate in accordance with their respective standards of rationality than to open up to the complex interplay of volitions and emotions that constitutes the moral practice. The problem is that in order to comprehend the moral sphere Tugendhat needs to refer to both the universalism that guides "communicative reason" and the considerations of prudence that are central to contractarianism. The two rival philosophers will claim that Tugendhat has to take the whole package - that is; he has to embrace a standard of rationality to the full. To them it is impossible to see morality as both a consequence

of self-interest and as governed by precepts of universalism. To describe morality as "symmetric contractarianism" is a contradiction in terms, they say.

But are these differences in method indicative of projects that are radically different in nature? That is; could it be claimed that Tugendhat simply has a descriptive aim while his critics strive towards something normative? This idea could seemingly find support in the fact that Habermas and Stemmer try to show how moral norms can be optimally rational, while Tugendhat clarifies how the phenomenon of morality should be understood. But this way of presenting the matter is begging the question about what normativity is. If one understands the concept of normativity as something exclusively rational then one could indeed contrast the differing accounts in this way. But since a central question of this thesis concerns the idea obligation it would be premature to use normativity as a classificatory notion. Both the idea of obligation and the concept of normativity rely on an understanding of the word "ought" (or rather the word "must", as we shall see). Tugendhat's claim is precisely that the normativity of moral obligation cannot be explained fully in terms of rationality.

In reality, the three different authors present answers to the same question: Given the nature of moral obligation, how can the norms be justified? But their different accounts of obligation lead to very different understandings of how norms can be justified. The impression one could get that Tugendhat's theory is "less normative" than the others has to do with his way of identifying what moral obligation is. His understanding of moral obligation is informed by a general concept of how the moral sphere should be understood. While the other authors presuppose that moral obligation must be of such a nature that no rational being can escape it, Tugendhat refuses this premise because of his understanding of the general anthropological basis of the moral practice. Before looking at how he envisions such an account we must look into the question of why a philosophical elucidation of the possibility of justification of moral norms is needed at all.

### *1.2 Modernity*

According to Tugendhat, the question of how we can justify our moral norms is pressing because we live in an age where people no longer know how to argue when they want to convince others that the moral norms are good: "Wir befinden uns heute in einer moralischer Desorientierung. Obwohl die meisten von uns durchaus

bestimmte moralische Überzeugungen haben, können sie doch gewöhnlich nicht sagen worauf sie beruhen" (Tugendhat 2006a, p. 13). He has always maintained that earlier times knew what was good, but "[w]ir haben heute diese Sicherheit verloren" (1979, p. 357).

What has changed today in comparison with earlier times, then? Implied in Tugendhat's description of the matter is that society has reached what Lawrence Kohlberg called the post-conventional stage, where norms are no longer perceived as handed to us by some external moral authority (Kohlberg 1981). The norms are not pre-arranged anymore, which means that it's up to us to take a stand on them and decide whether we can accept them. The problem is that we do not know how to justify any criteria for deciding when a norm is acceptable or not. We have many opinions that we do not know how to justify without being dogmatic. Our moral judgements have an origin in a religious past, but since religious justifications no longer persuade we do not know what we can appeal to.

There is an ambiguity here: Whom does Tugendhat refer to by "we"? Although he has written much about whom the "we" of the modern moral community should be - meaning whom we owe moral respect - he has written nothing about whom he has in mind when he refers to people in a state of disorientation. Presumably he means those that do not believe in any "higher authority" anymore. This is plausible because he sometimes says he is investigating whether there can be a "modern concept of morality". The absence of a transcendent source of moral legitimacy is the defining feature of Tugendhat's notion of modernity: "Unter moderner Moral verstehe ich einfach eine Moral, die sich nicht auf die Autorität von Religion und Tradition stützt" (1996, p. 323).

This means that modernity is understood as a conceptual notion, not an historical epoch. This way of using the notion of modernity allows him to describe Protagoras of ancient Greece as an early representative of a modern moral consciousness (cf. *ibid.*, 327). It is enough to reject an external and "higher" moral authority to be modern in Tugendhat's sense; one does not have to be born in any particular century. The "higher" authority could be some divinity, a legend or tradition. The important aspect is that norms that are given by a "higher" authority presuppose a belief in the moral supremacy of this being or tradition. So by casually referring to "we", Tugendhat is referring to those that do not believe that matters of

right and wrong can be answered by appeal to something higher than our ordinary interests. But the question is how far one gets by referring to "us moderns" in this sense, then.

First we have to know what Tugendhat wants to achieve by starting off with this claim about our present situation. It is a descriptive claim, referring to an actual confusion that arises when people try to justify moral opinions. The philosophical problem that he pursues is supposed to start off from a real sense of uncertainty in our moral belief system. But the problem with this empirical claim is that we do not know its extension. Moral philosophers have stopped anchoring their argumentation in theology a long time ago, but Tugendhat presumably wants to cover a more significant group of agents than these. He seems to say that people in general have lost faith in traditional and religious justifications, and that this is what creates a problem in our interaction.

Is he saying that even though some people still believe in some higher authority we must assume this kind of justification to be subsiding? That this kind of justification is losing its grip in society? This is in accordance with the well-known social hypothesis of increasing rationalisation and secularisation. Science is thought to gradually replace irrational religious beliefs. Modernity is tied to a process of disenchantment, Max Weber claimed.

Nevertheless, Habermas now speaks of a post-secular age, suggesting that the theory of ever more secularisation is being falsified before our eyes: "Die lange Zeit unbestrittene These, dass zwischen der Modernisierung der Gesellschaft und der Säkularisierung der Bevölkerung ein enger Zusammenhang besteht, findet unter Soziologen immer weniger Anhänger" (2008, p. 34). He lists several factors that should make us doubt that religion has loosened its grip in society, and advises us to deal with this as a feature of the present age and not the past.

Now Tugendhat may be right in his description, but he may just as well be wrong. That is an empirical question we cannot settle here. But does Tugendhat *have to* figure this out in order to get his theory off the ground? Perhaps not, since he could for instance go along with Rawls and state that "a basic feature of democracy is the fact of reasonable pluralism - the fact that a plurality of conflicting reasonable comprehensive doctrines, religious, philosophical and moral, is the normal result of its culture of free institutions" (Rawls 1997:441) That is, the fact of *pluralism* and not

primarily disorientation on the part of the agents creates the need for a modern concept of morality.

How does this help? It relocates the problem from the subjects to the intersubjective. This way of seeing things comprehends the disorientation as a result of the encounter between people of a radically different worldviews, and does not predicate any confusion on the parts of the individual agents or various particular moral communities. This avoids the impression that the philosopher has wished away the fact that people still have firm moral judgements which they can fully justify, at least to those that share their beliefs. The problem is that they may often be incapable of justifying their views to other groups. Since many moral conflicts both transcend and yet strongly affect particular doctrinal groups it is necessary to build upon premises they can mutually accept. This means that any concept of morality that is to be plausible in a pluralistic society must be disconnected from any particular "higher truth". The point is, however, that we cannot expect this common justification to replace the need for a final and personal justification in something transcendent. The modern premises may not go "all the way down", so to speak. The moral norms may perhaps only be backed by what Rawls calls an "overlapping consensus".

So it seems like incorporating the ideas of reasonable pluralism and overlapping consensus is a solution to Tugendhat's ambiguous description of the problem. Unfortunately, however, these "political" concepts put an end to the whole affair of justifying a moral standard. Overlapping consensus is tied to the idea of a "freestanding view of justice", which can consolidate "reasonable" religious, philosophical and moral doctrines (cf. Rawls 1993, p. 12ff.). This implies renouncing the attempt to give universalism a philosophical grounding, since the justification is rather handed to the agents themselves. Instead of justifying the universalism that underlies the idea of a liberal society, Rawls simply starts "from within a certain political tradition" (ibid., 14).

So there is a problem in using the idea of overlapping consensus, since it reduces moral justification to the quest for some *module* that is neutral with regards to the various moral belief systems. The idea that there is a pluralism of doctrines towards which philosophy must be tolerant is not compatible with the project of justifying moral obligation independent of any "higher authorities". So this cannot be the starting point for Tugendhat, since he does not want to construct an argument

"independent of controversial philosophical and religious doctrines" (Rawls 1985, p. 223). He wants to give an account of how the principles that the liberal tradition relies on should be justified and understood, and not simply "start from within" this tradition.

His project therefore resembles Kant's *Grundlegung* more than anything. Kant humbly formulated his aim as "nichts mehr als die Aufsuchung und Festsetzung *des obersten Prinzips der Moralität*" (1785, p. 392). This project neither needs nor should be grounded in some disorientation found in society. The only way to perceive this societal confusion as a problem is by making the judgement that we should be able to justify our behaviour to each other. But this idea of reciprocal justification actually relies on a notion of universalism, that is, that we are to give reasons for our actions instead of using force. This is the very principle that is to be explained, it is therefore unfortunate to use it in attempting to make a neutral observation of a problem.

In order to avoid this whole problem we must read Tugendhat's claim differently (and perhaps divergent from his intentions). It is not the general "we" of society that is in a state of confusion, but those of us that want to justify our moral opinions by appeal to our common interests. From now on, this is what "we" means. This implies that we are no longer to start from a disorientation found among the members of a community, but ask ourselves whether we can make sense of the moral practice without presupposing a belief in some moral authority. Are terms like 'obligation' and the standard of universalism justifiable features of a secular morality?

Why should we care about this, if the motive no longer stems from a societal sense of disorientation? We should care because the philosopher can help identify and explain how moral norms that people may or may not perceive as given by some authority can be understood even independent of a transcendent basis. Furthermore, the fact that people do not find themselves in a state of disorientation does not mean that their opinions may not be confused. Consider those who follow Nietzsche in saying that moral obligation died with God. The philosopher can clarify whether this relies on a faulty understanding of the basis of our moral norms or not. The philosopher can identify how moral norms are situated beyond both religion and scepticism. Without this elucidation the norms are surrendered to a precarious existence. As Kant stated; "die Sitten selber [bleiben] allerlei Verderbnis unterworfen,

solange jener Leitfaden und oberste Norm ihrer richtigen Beurteilung fehlt" (1785:390).

Of course, we do not yet know whether such a justification is possible. Perhaps the moral standards are mere leftovers of a religious tradition and Nietzsche was right in saying that to live a life in truth and to live one regulated by moral principles is incompatible. It could be that morality is the result of a human need to create helpful fictions (cf. Nietzsche 1885, § 4). In the following I will investigate how Tugendhat faces these speculations and demonstrates that morality can be understood independently of all higher authorities. We should imagine him as less tormented version of Ivan in Dostoyevsky's *The Brothers Karamazov*, who asked in despair: "But what will become of men then?" I asked him, 'without God and immortal life? All things are lawful then, they can do what they like?' (Part IV, book 11)



## 2. *What is Morality?*

As was mentioned in the previous chapter, Tugendhat aims at understanding our moral obligation to the standard universalism in light of a general account of what morality is. The problem is how moral norms can be understood in the absence of some transcendent moral authority, and in order to see how moral norms can be justified at all he needs a concept of morality that allows him to compare both traditional and modern justifications. What is common to both the traditional and the modern uses of concepts like 'obligation' and 'good'?

A major part of his writings on moral philosophy consists in establishing such a description of the practice itself. In fact, most of what he has to say on the justification of moral standards hinges on this "formale Vorbetrachtung" on morality. Giving a basic account of the concept of morality means searching for the necessary features of what we call moral practices. This must be of such a quality that it allows for both a comparison of specific moral accounts and a separation of the moral practice from other kinds of social interaction.

The description of the moral practice aims at neutrality in order to avoid predetermining any genuine disagreements by the mere understanding of the meaning of a word: "Wir müssen das wort 'Moral' so weit definieren, daß wir verschiedene Moralkonzepte unterscheiden und miteindander vergleichen können" (Tugendhat 1993, p. 33). Now, it will become clear that this formal concept of morality is not neutral in the sense that this quote could seem to imply: Morality is not defined so loosely that any moral standard could be foundational. The standard of universalism is there from the start, although Tugendhat has only recently made this explicit. But the formal concept is neutral in the sense of being compatible with both justifications that appeal to some "higher authority" and justifications that refer to our ordinary interests.

Furthermore, we must be clear about the objective of gaining this formal concept of morality. Tugendhat does not want to discover some "true" meaning of the word, as he has no belief in such nominal essentialism; "es ist nie sinnvoll, sich über den wahren Sinn eines Wortes zu streiten" (1999a, p. 165). What matters is gaining an understanding that allows us to compare different configurations of morality, and to understand what moral justification can amount to. So this formal concept of morality must be evaluated according to its pragmatic use.

This chapter starts with a presentation of Tugendhat's ideas on what moral obligation is (2.1). It proceeds to a discussion of how he understands the moral use of 'good' (2.2). This understanding of 'good' relies on a more general theory of human cooperation, which will be presented in chapter 3. What is at issue in the present chapter is first and foremost the recognition of how moral obligation stands in relation to moral goodness. So this chapter discusses a certain conceptual interrelation, the necessity of which will be explained in light of Tugendhat's thesis of "collective actions" in the next chapter. The final section of this chapter deals briefly with Dieter Birnbacher's critique of Tugendhat's concept of moral obligation. I mention this criticism because it has an intuitive appeal that must be account of.

### *2.1 Moral Obligation*

How does one proceed when attempting to achieve a formal description of the moral practice as such? Obviously there are certain criteria to be met, not any formal concept that encompasses both traditional and modern justifications will do. Studying human practices must strive to achieve what Anthony Giddens called a "double hermeneutic" (Giddens 1976). We are trying to understand interaction among people, and these people have their own understanding of this interaction. This means that the concepts that the theorists use must have a clear relation to the concepts that the subjects use themselves. Our understanding must be based on their self-understanding: "Social life cannot even be accurately described by a sociological observer, let alone causally elucidated, if that observer does not master the array of concepts employed (discursively or non-discursively) by those involved" (Giddens 1987, p. 18f.).

Does this methodological stance matter to philosophy? It may seem like it doesn't, since it deals with clarification of our basic concepts and attitudes, which is something else than finding out empirically how people understand themselves. But this project of clarification cannot be abstracted completely from the actual language games. We cannot understand concepts without looking at how they are used in everyday interaction, meaning that we need to know how the moral concepts *function*. This gives us a strong incentive to pay attention to Tugendhat's semantic method of capturing the defining features of morality. We express ourselves linguistically, so the raw material for an understanding of morality is found in the language people use.

Tugendhat first narrows down some key words commonly used in moral discourse; "*good*", "*bad*", "*must*" and "*can't*" (cf. 1993, p. 36f.). The former two are referred to as evaluative expressions (*Wertausdrücke*), while the latter ones are what he calls necessity expressions (*Notwendigkeitsausdrücke*). "Good" is the topic of the next section, and we will see that it stands in an important relation to "must", which is the issue of the current section. We cannot have one of these moral expressions without the other.

The goal of finding an explanation of the necessity referred to in moral contexts ("You must respect others!") is of primary importance to Tugendhat. This is revealed with all possible clarity in a statement that is programmatic of his whole moral philosophy: "Die Größte Gefahr der Verunklärung, die der Moralphilosophie droht, ist, den Sinn des 'soll' in einem Nebel zu belassen" (1981a, p. 73). His notion of moral obligation is one of the most central points in his formal concept morality.

The first encounter we have with morality in our upbringing is that certain actions are allowed while others aren't. We are told that there are general rules that everybody *has to* respect. To gain an understanding of the moral sphere we must know what it *means* for an action to be forbidden before we ask *why* it is forbidden. What does the moral "must" mean? First a word on why one should speak of "must" instead of "ought". Tugendhat claims that the usual orientation in moral philosophy towards the word "ought" (*soll*) is unfortunate, since it does not convey the categorical nature of moral demands: "Man soll nicht nur sein Versprechen halten, man muß es" (cf. *ibid.*).<sup>1</sup> This is in agreement with J.L. Mackie, who demonstrated the awkwardness of using "ought" by inserting it into the Ten Commandment instead of "shall": "'You ought not have any other Gods before me ... You ought not kill'" (1977, p. 64). This drastically changes the mood the Old Testament and reminds us that "ought" is used when we advise people on what they should do. But we don't advise people to be moral. Moral behaviour is *demande*d of us. What is central to the moral vocabulary that Tugendhat has identified is that it presents each agent with a *practical necessity*; some behaviour is morally *good* and therefore it *must* be respected.

But what constitutes this necessity? What disables us to be immoral? Surely it's not physically impossible, since we would not take it to be a miracle if someone

---

<sup>1</sup> Peter Stemmer offers an extensive account of how it came about that philosophy after Kant deals with "ought" and not "must" (cf. 2008, p. 289ff.).

broke his or her promise, for instance. And we aren't *forced* by others to be moral either. This would have made morality paradoxical, since forcing others to do something they do not want to do is usually considered immoral. Moral necessity is that enigmatic force called normativity. This force does not reign in the same sphere as the laws of nature. As Stemmer succinctly puts it: "Die Naturgesetze wirken gewissermaßen an unserem Kopf vorbei, das normative Müssen nur durch unseren Kopf hindurch" (2008, p. 55). It seems to some extent to be "up to us" whether we want to be moral. But at the same time the moral demands are considered *unconditional*. How is this possible?

Tugendhat finds only one previous answer to this question, and he takes it to be wildly implausible. Kant also believed that the normative laws only work through our consciousness, but he imagined this relationship in a very special way: We are members of two different worlds. We are both natural and transcendental subjects. He claimed that the moral law expressed in the categorical imperative was issued from a transcendental "pure reason", and to the extent we want to be rational we must act in accordance with this law. But this is far too heavy metaphysics for Tugendhat to accept; "es liegt nahe, Schopenhauer darin zuzustimmen, daß die Idee dieser angeblichen reinen praktischen Vernunft ein Versuch war, die Vorstellung eines religiös fundierten moralischen Gebotes zu säkularisieren" (Tugendhat 2006a, p. 15).

In order to close in on an adequate understanding of moral obligation Tugendhat goes to the social sciences. This is a natural move, since the concept he wants to operate with is supposed to be one that an ethnologist could use as well. Tugendhat borrows a common sociological description of what is considered moral behaviour: "Regelmäßigkeiten im Verhalten, die auf sozialem Druck beruhen" (Tugendhat 1999a, p. 163). What is gained by this definition? The reference to "social pressure" reveals that the force of normativity is something social. But this is too unspecific if one wants to distinguish the moral norms from other kinds of social norms. What is specific about the pressure connected to the moral norms are the kinds of emotions that are involved and their reference to an objective standard. So in order to explain the way our moral obligation is constituted in the social sphere Tugendhat asks us to consider what happens to the person that transgresses the moral norms: "wenn ein Mitglied gegen die gemeinsam akzeptierten Normen verstößt, [reagieren]

die anderen mit einem negativen Affekt, und man kann diesen Affekt als Empörung oder Entrüstung bezeichnen" (Tugendhat 2006a, p. 19).

This statement contains one of the most provocative elements in Tugendhat's moral theory. He claims that the moral obligation is *constituted* through the emotional sanctions that occur when norms are transgressed. Why is this considered provocative? That is mainly because moral obligation is usually taken to be something *prior* to sanctions. We sanction each other *because* we are all under a moral obligation. People are supposed to do the right thing because it is right, not because others will react negatively if they don't. This critique is important, but will be postponed to the end of this chapter. It is crucial to get a better understanding of what Tugendhat actually means first. To speak of sanctions in this loose fashion gives the wrong impression of what is implied in his theory.

As opposed to the "external" and physical punitive measures taken in the legal sphere, the moral sanctions are "inner" and emotional reactions connected to a moral judgement.<sup>2</sup> Tugendhat draws on the seminal essay "Freedom and Resentment" (1962) by Peter Strawson. This deals, among other things, with the propositional content of our reactive attitudes or moral sentiments. Tugendhat therefore speaks of the "Strawsonian Triad" of *indignation*, *resentment* and *guilt*: "Entrüstung, Groll und Schuldgefühl; in diesen negativen Gefühlen konstituere sich erst das moralische Sollen" (Tugendhat 1997a:11). It is important to notice that both Strawson and Tugendhat have a very specific idea about the content of these emotions, and the way they use the terms may vary from everyday use. They are technical terms that indicate idealised reactive patterns. Our actual reactions are usually messier than what these terms convey, varying both in spontaneity and content.

The Strawsonian Triad refers to three different perspectives we may have on behaviour. *Indignation* is a "vicarious" attitude, in that we judge that some conduct that didn't concern us directly was bad. We react only as bystanders to an act of ill will. When we ourselves are the victims we react with *resentment*. And the agent who commits the wrongdoing is prone to feel *guilt* as a consequence of having transgressed a norm that he takes to be justified. Strawson argued that moral reactive

---

<sup>2</sup> Tugendhat avoids the using words like "inner" and "outer" sanctions in his later writings, most likely because they are too different in nature to be distinguished by these spatial metaphors. We, on the other hand, should stick to this terminology, since in 5.1 it will be clear that he has maintained a dubious analogy between the two groups in spite of his rejection of spatial terminology.

attitudes are connected in such a way that if you react with resentment you must also be susceptible to guilt. The same goes for the vicarious attitude of indignation: To the extent attitudes of disapproval of *others* manifest themselves in us we must also be open to *self-disapprobation* when we transgress the same norms. If I only blame others for their transgressions I would illustrate "an abnormal case of moral egocentricity", according to Strawson, and this is something he considers "barely more than a conceptual possibility: if it is that" (Strawson 1962, p. 85).

Tugendhat seems to agree with this, but one could argue that Strawson exaggerates the logical consistency of our emotional apparatus. Surely we may occasionally encounter the self-righteous person who is deliberately blind to his own failings yet never misses an opportunity to censure others. But then again, it is doubtful that his resentment of others would be completely sincere. Nobody who wholeheartedly identifies with the moral norms shows such deliberate blindness towards own transgressions. Nevertheless, it can't be denied that some people are fully unsusceptible to moral sentiments, even though these are rarely encountered. In various contexts Tugendhat mentions the case in which a person has "lack of moral sense". This may be the result of an unsuccessful socialisation (or perhaps too much Nietzsche).

Importantly, the moral "must" cannot apply to an agent who does not have the emotions that were identified in the Strawsonian Triad: "Sie kann nur diejenigen treffen, die sich als zugehörig zur moralischen Gemeinschaft verstehen" (1989b, p. 322). This brings out another important aspect of Tugendhat's theory: To be a member of the moral community is to have the moral sentiments connected to the norms of this community. This means that people can be immoral in two ways: Either by not belonging to a moral community (lack of moral sense), or by being a member and yet prioritise personal benefits above a clean conscience (cf. Tugendhat 2006a, p. 22). In the latter instance we can say that the agent considers his prudential good of higher importance than the moral good. But what does "moral goodness" mean?

## 2.2 Moral Goodness

Tugendhat claims that "must" and "good" are used equivalently in moral contexts, in that they both are used in a grammatically absolute way and present the agent with the practical necessity of conforming to a the norm (cf. 1993, p. 36). What is meant is that expressions like "You can't do that!" and "That's bad" are the same perlocutionary act.

However, they do not *mean* exactly the same (different illocutionary acts<sup>3</sup>), in that "good" is an evaluative expression that necessarily refers to some standard.

A moral conscience must be constituted in relation to some objective principle - it cannot revolve around what the agent admits to be merely subjective preferences. Tugendhat's point is that the phenomenon that has been identified as moral obligation cannot be explained in terms of mere self-interest: Moral obligation would be impossible because the moral sentiments that grounded the concept of moral obligation have a propositional content that cannot be reduced to some instrumental calculation relative to pre-moral interests. We cannot resent people without claiming that they transgress some *standard* that all are supposed to adhere to. We may get angry or frustrated without having accepted some objective principle, but in order to experience resentment and guilt we need to refer to a standard and not mere self-interest.

So far we have seen the negative story: We cannot keep our moral sentiments without some idea of the morally good. But what is the positive story? That is, how does one make sense of a moral use of "good" that is supposed to regulate our pursuit of personal interests? The appeal to the propositional content of our sentiments is no proof that one can make sense of an objective standard: "The claim to objectivity, however ingrained in our language and thought, is not self-validating" (Mackie, p. 1977). People may refer to an allegedly objective basis when they argue morally, but that does not mean that it exists. Perhaps we have such strong *need* for a solid basis for our moral ideals that we have retained some metaphysical projections in our language to meet this

Tugendhat's task is to show that his suspicion is mistaken. In order to understand the moral use of "good" he looks at practical evaluations in general. Every performance of an activity can be evaluated along a scale of better and worse. Usually this standard is intersubjectively recognised, in that there is some common and agreed upon way to separate the good violinist or chef from the bad. Tugendhat calls such standards of evaluation "adverbial" and claims that "das moralisch Gute ein - freilich besonderer - Fall des adverbialen Guten ist" (2003, p. 70). What makes the moral standard of evaluation special is that one is not good in some specific regard. Rather,

---

<sup>3</sup> Tugendhat does not use J.L Austin's terminology of illocutionary and perlocutionary speech acts, but this distinction is implicit in his talk of an equivalent use of different terms that refer to "verschiedenen Nuancen der moralischen Urteil" (1993, p. 37).

one is simply deemed good or bad as a person. But who is to decide what makes a person good?

The members of the community must decide this for themselves. Tugendhat therefore borrows a definition of the morally good person from Rawls: The morally good person has "the features of moral character that it is rational for members of a well-ordered society to want in their associates" (Rawls 384<sup>4</sup>, §66). Here "rational" only means in our interest, so moral goodness refers to our mutual wants. This way of understanding moral goodness implies a notion that is of central importance to Tugendhat, namely *reciprocity*. "Good" as defined by Rawls refers to our shared interests, and is a formal description of the intersubjective and reciprocal nature of moral goodness. It does not refer to some concrete function, like being a good chef refers to making good food, but only to whatever qualities people want *in each other*. So the moral standard, as opposed to other adverbial standards, refers to the way *all* should be. Only a few people are expected to be good chefs, but each member of our social community is expected to be moral. A moral imperative is therefore bi-directional; morally demanding something of others signals the will to demand the same of oneself.

The criterion of reciprocity gives us an idea of what is meant when moral norms are said to be unconditional. The categorical nature of moral imperatives means that this sort of goodness is expected no matter what specialised talents we want to develop or what position we hold in society (cf. Tugendhat 1997a, p. 37). This means that the idea of a "categorical" norm is understood differently than how Kant interpreted it. It does not mean that there is an unconditional practical necessity for each rational being to be moral, but that there is a social expectation of a certain kind of behaviour *independent of our personal preferences and talents*. This expectation is not grounded in rationality, but in reciprocal interests.

Rawls' definition of the morally good person has direct consequences for the understanding of a morally good norm. Since all are equally expected to adhere to the norms they must be *equally good for all* (cf. Tugendhat 1998, p. 101). If all are to be restrained to the same degree the norms have to be in the interest of all to the same degree as well. If the norms are better for some than others this means other considerations than what can be morally justified have decided the outcome.

---

<sup>4</sup> Page number refers to revised edition, 1999. Tugendhat uses this definition in 1997, p. 37 and subsequent writings.



But surely one may find empirical evidence that moral norms have not always been equally good for all. Tugendhat does not need to deny this, all he wants to show is that this inequality is based on non-moral premises. Norms of inequality could for instance evolve if a section of the community had the means to force such norms through, in which case no justification is needed. Force replaces moral justification. This will be further discussed in light of Stemmer's contractarian account of legitimate norms, where a moral norm merely does away with the need for an actual battle (chapter 7). Tugendhat's claim is morality cannot be understood in this non-universal way without losing the feature of intersubjective obligation.

Another cause of moral inequality is when there is a unified belief in some "higher authority" that decides that some should have rights that others do not have. But in this case equal *faith* replaces the equality among interests (cf. Tugendhat 1997, p. 75). So the principle of universalism is constitutive of both traditional and modern moral practices. People are not forced to believe in the "higher authority", and so the inegalitarian norms issued from this source could be morally justified.

Justified moral norms are thereby *defined* to promote the idea of non-discrimination and non-coerciveness. The critical question presents itself: Why do we have to understand moral norms and goodness in this way? This way of seeing morality implies a standard of reciprocal and equal recognition of all, but how can Tugendhat just postulate that norms must satisfy this criterion? As suggested above, Tugendhat is trying to define morality as the opposite of coercion. The problem is to get beyond mere definitions and down to the human framework of actions and attitudes that underlie these. This is the issue of the next chapter, where we will see how Tugendhat relates universalism to the preconditions of non-coercive joint action. This means interpreting moral norms in light of the general anthropological basis of cooperation. But for now we must try to understand what it means that all are equally restrained by morality.

This can be elucidated with help of Tugendhat's concept of "collective autonomy" (cf. Tugendhat 1998, p. 98f.). His idea of moral autonomy does not consist in every member of the moral community giving him- or herself the moral law, like Kant envisioned. The "collective autonomy" is rather that adherence to the principle of equal respect is mutually demanded. This is the meaning of his concept of reciprocity; people are equally restrained by the moral norms since they are equally

demanded of all. Such a view of moral autonomy can be illustrated by imagining the moral practice as a special kind of chain gang, where each prisoner holds a key to his ankle shackle. The problem is that unlocking one ankle shackle implies unlocking all. The prisoners could cooperate as long as all were chained, but when let loose they fall prey to each other. The possibility of the moral practice depends of people's willingness to shackle up, and nobody would do that unless the rest did so as well. Importantly, the prisoners are free as a collective; the shackles do not come from some external source since each member holds a key. And each member is only restrained to the extent that he wants all to be.

So when we transgress some moral norm we are deemed bad according to this special intersubjective standard of evaluating behaviour: The propositional content of our reactive attitudes refer to the standard of reciprocity and collective autonomy. The attitudes of resentment and guilt contain a judgement that the transgressor's behaviour is not compatible the mutual interests in his community. He has made himself an exception and put his own interests above those of others. It is this feature of immorality that makes resentment and indignation into something non-personal, since those that perceive this transgression do not only react with frustration on their own behalf, but think that the *whole* moral community should condemn this behaviour.

### *2.3 Emotions of Self-Assessment and Obligation*

It is not necessary to discuss in any further detail whether his account of our moral sentiments is comprehensive or in tune with our everyday experience. There may certainly be other descriptions that have greater phenomenological appeal, and what has been presented so far is perhaps too abstract to be an account of how these emotions are experienced. But Tugendhat has never aspired to write any general phenomenology of moral sentiments. He has rather attempted to prove that they are the basis of moral obligation. His account has at started off from something tangible that we can relate to; we all know that pangs of conscience is something we would rather avoid. It is in the nature of guilt to warn us that we have done something *alien to our self-conception*. If one wants to attack Tugendhat's account of the moral sentiments, it has to take the shape of an accusation that he has failed to explain the concept of moral obligation. Has he failed this?

Many critics have found the idea of tying moral obligation to inner sanctions counterintuitive. The common accusation is that Tugendhat concedes too much to pre-

moral concepts of rationality by describing moral behaviour in this way. Dieter Birnbacher claims that Tugendhat mistakenly reduces moral behaviour to self-interested calculation: "Nicht die Angst vor Schuldgefühlen lässt uns vor dem moralisch Verwerflichen zurückweichen, sondern die Ablehnung des Verwerflichen selbst" (2003, p. 326). The idea is that moral behaviour must rather be explained in terms of what we perceive as good and bad. That some kind of behaviour appears as morally bad is enough to explain the moral obligation we have to avoid it. We do not need recourse to the idea of "inner sanctions" at all.

Firstly: It is a grave misrepresentation of Tugendhat's position to say that moral behaviour is reduced to strategic calculation, as Birnbacher does. This overlooks the propositional content of moral guilt, namely that one acted in a way that is acknowledged as *bad*, not just undesirable. This presupposes a conscience and a self-understanding as somebody who does not only calculate instrumentally according to selfish interests. Secondly: Tugendhat has actually *given content* to the idea of "doing something because it appears morally good". The attitudes of guilt and resentment are used as a substantial alternative to using empty concepts like "Ablehnung des Verwerflichen selbst". To explain our obligation in terms of emotional attachment to a moral standard is not contrary to saying that it is the morally reprehensible that we reject. It is rather a way of unpacking our standard ways of speaking, showing how it relates to our basic attitudes.

The emptiness of Birnbacher's criticism is even more evident in the preceding sentence: "Wenn es uns wirklich wichtig ist, kalkulieren wir nicht" (ibid.). This amounts to the tautological claim that we don't weigh our prudential good against the moral good *if they overlap*. He is getting at the experience we sometimes have of knowing what we have to do without any deliberation. But this is not the interesting case. What is philosophically important is figuring out what it means to say that people have to behave morally even when it is contrary to their prudential good. This question only makes sense if their personal desires are not already shaped completely in line with moral norms. In what sense do they *have to* choose the moral alternative? Saying that we often act morally without considering what the consequences of immoral behaviour would be is not contrary to Tugendhat's account.

In order to criticise Tugendhat's account of the moral "must" one should confront him with an alternative account. According to Tugendhat, only Kant that had

such an alternative account. As will be seen in Part 2, Habermas is one of today's most vigilant defenders of the Kantian "Achtung vor dem Gesetz". He claims that this dictum contains account of moral obligation which is superior to the sanction-based. But before we can understand Habermas' criticism we need to examine how Tugendhat himself thinks that the Kantian universalism that is implied in his concept of morality is to be understood. So far I presented the way he stipulates certain descriptions of moral goodness. It has been shown that his definition of morality takes him towards an understanding of moral goodness in terms of equal respect for all. But the question is why morality has to be understood in this sense. Why couldn't morality be based on a different standard or no principle at all?

### 3. *Why Universalism?*

In this chapter I will discuss Tugendhat's explanation of why moral goodness needs to be understood in terms of reciprocal respect for all. We are looking for the rationale behind treating "collective autonomy" (that is achieved in putting each other under the same restrictions) as the criterion of good moral norms (cf. 2.2). Couldn't moral norms be upheld in society in order to satisfy a different goal? The idea that norms must be based in our reciprocal interests is not the only way to understand them.

They could alternatively be said to promote the greatest happiness or the private interests of each member of society. That is; why shouldn't morality be understood in a utilitarian or a Hobbesian fashion? Contractarianism in the Hobbesian sense will be discussed in chapter 7, where Peter Stemmer's criticism of Tugendhat will be considered. For now it suffices to say that Tugendhat considers this kind of contractarianism the only plausible *starting point*. Morality needs to be based in the ordinary interests of the members of the moral community, but Tugendhat thinks that the criterion of universalism must be added to this story: Norms must be in the equal interests of all, not just those that can contribute to our own benefit.

With utilitarianism the case is different. Tugendhat does not consider the greatest happiness principle to be more than a bi-product of capitalism (cf. 1993, p. 327). It is not my business to discuss utilitarianism at any length here, but a comparison is useful in order to see what universalism actually implies. In utilitarian theories everybody is considered equal, but in a strictly numerical sense: "Everybody to count for one, nobody for more than one" (ascribed to Bentham by J.S. Mill, 1861, p. 64). Just as two identical coins have equal worth so do humans. This sort of worth is not the moral respect that universalism prescribes. It can be outweighed by quantitative measures; we gladly exchange ten coins for twenty of the same kind. Similarly, if a large group of people benefited greatly from the repression of a few it would be morally allowed to repress them according to a strict utilitarian principle. Some, like John Stuart Mill, built safeguards into the utilitarian theory in order to prevent this. But this cannot be made sense of on utilitarian grounds alone. As Isaiah Berlin pointed out, Mill's deepest moral convictions are not contained in his formal arguments (cf. 1969, p. 200f.).

The question is why one should build safeguards against tyranny of the majority at all. What is wrong with torturing a few to protect the citizens of a country,

for instance? In other words: Why is the moral ideal of mutual respect morally good, while advancing the greatest total benefit at the cost of a few not? Stating an opposition between the principles like this might seem wrong. Some claim that we need to recognise that there is not *one* highest principle - there are several and they are legitimate within different spheres. Tugendhat would deny this: Universalism as understood in light of the categorical imperative is the only fundamental principle. When using the utilitarian principle to decide a strategy of action this must be equally justifiable among the members of the moral community (I will return to this point in 4.2). It could for instance be in the equal interest of everyone that lifeguards act according to the rule of saving large groups instead of sole swimmers when this dilemma occurs. The principle of maximisation can be a means to satisfy the standard of universalism then, but not a moral standard on its own. How can Tugendhat defend this without already presupposing the principle of universalism, then?

Since Tugendhat is trying to construct a formal concept of morality at this stage he cannot rely on our current moral consciousness. He cannot simply claim that the greatest happiness principle "does not seem sufficiently closely linked to our ideas of right and wrong" (Scanlon 1998, p. 152). If Tugendhat manages to make it plausible that any moral practice must be understood according to the standard of universalism he has an argument against utilitarianism that does not rely on what we happen to think is right today in our Western society.

This chapter is divided into four sections in order to properly identify each step in the account of universalism: 3.1 briefly comments upon the way Tugendhat's position has changed when it comes to giving an account of why morality demands equal respect of all. 3.2 is a presentation of the way he uses two quotes from Aristotle in order to illuminate the specifically human preconditions for cooperation. In 3.3 we will see how he explains the notion of universalism as one of two fundamental ways of structuring a cooperative action. 3.4 is an explanation of how Tugendhat envisions that the alternative of universalism gains normative force.

### *3.1 Pioneering Investigation*

The understanding of universalism that Tugendhat finally arrived at was presented in the essay called "Der Ursprung der Gleichheit in Recht und Moral" (2007b). The title conveys the ambitious nature of this enterprise. In his previous attempt he noted that such an undertaking is missing in the whole history of ethics. He claims to have

"keine Vorgänger" in the effort to answer this question without being dogmatic or relying merely on fault-finding in competing accounts (cf. 1997a, p. 68). But as will be explained, his answer is far from plucked from out of mid-air - in spite of the claim to novelty. In the presentation of Tugendhat's solution I will be looking closely at his use of Aristotle and Habermas, since his final solution is a fusion of ideas from both of them.

A difficulty for Tugendhat when it comes to demonstrating why universalism has a place in morality is that his toolbox is rather scanty. He has neither Kant's transcendental reason nor Habermas' communicative reason at his disposal. Rawls and many others that promote the idea of equal respect for all consider it unnecessary to demonstrate the basis of the principle itself. "Reflective equilibrium" steps in as a substitute of philosophical elucidation. This is fully in accordance with the idea that there is "no noncircular theoretical backup for the belief that cruelty is horrible" (Rorty 1989, p. xv). Richard Rorty achieved his status of philosophical *enfant terrible* simply by being somewhat blunter than his many of his colleagues.

Tugendhat's solution involves looking at two alternative ways of structuring human interaction. It is worth emphasising the role Tugendhat has given to philosophical anthropology in his later writings. It contrasts the central position he gave semantics in his early efforts. Locating the origin of universalism in the structure of human cooperation is miles away from his solution in the essay "Sprache und Ethik": There he briefly noted that moral judgements stand in some relation to basic anthropological features of social cooperation, but believed that a justification of moral goodness could be achieved independent of these other phenomena (cf. 1981b, p. 291). "Good" was thought to have a specific moral use that was distinct from all other uses. By the end of his quest he has come to believe the opposite; moral goodness is an adverbial evaluation that is similar to the evaluation of other activities, and it is based in a specifically anthropological way of cooperation.

Tugendhat's focus on philosophical anthropology invites to a more fundamental perspective on human behaviour than the usual orientation towards interaction that is already regulated by moral principles. If universalism is a necessary ingredient in the moral practice as such, then the explanation of this principle should lie in the features that enable humans to engage in this sort of practice in the first place. Thoughts like these seem to underlie his final solution to the question of the

origin of universalism. And this makes it understandable why he has reconsidered Aristotle's views on the way moral communities come into being.

### 3.2 Agreement as Pre-Condition for Collective Actions

There are two related ideas from Aristotle that are important here; the first is the nature of our will and the second is the way this will can be put in relation to other wills. According to Tugendhat, the specific trait of us humans is the fact that we have a language that allows us to construct propositions (cf. 2007a:42). He refers to the famous passage in Aristotle's *Politics* (1253a10) where it is said that man is the only creature endowed with *lógos* (here meaning a propositional language). Our kind of language can release us from the impulse-governed nature of animal languages. We can act on *reasons* instead of just responding to our immediate reactions.<sup>5</sup> Instead of just reacting to an impulse we can take a stand on how to be: We can consider alternatives and make priorities. But how is the ability to construct propositions essential to the ability to prioritise according to what we think is good?

"Eine Vorstellung von Gutem könne man im Gegensatz zu einem Gefühl von Angenehmen nur Haben, wenn man es als Prädikat versteht" (Tugendhat 2003:14). So the predicative language is a necessary condition for the use of the concept of 'good'. We must be able to capture the judgement in a sentence that we can say yes or no to, which presupposes the ability to decide whether a general term is true of the singular term.<sup>6</sup> This gives us the opportunity to ask whether this option is *better* than that, and the freedom to decide what to pursue on account of our evaluation of options. And importantly; we can communicate our understanding of what is good to others. We can discuss what is worth pursuing and how to pursue it, and in this way we can agree to a common goal. If others agree to our understanding of what should be done we could *cooperate*. We could perform what Tugendhat calls a "gemeinsame Handlung" (cf. 2007b, p. 144).

---

<sup>5</sup> Tugendhat's talk of reasons is a small mystery. In his later writings he refers to Scanlon's account of reasons (cf. Scanlon 1998, chapter 1). But this is part of an externalist thesis that denies any inherent connection between the desires and practical rationality: [I]t is almost never the case that a person has a reason to do something because it would satisfy a desire that he or she has" (ibid., 8). Reasons exist in a different realm than our desires; "being a reason for' is an unanalyzable, normative, hence non-natural relation" (ibid., 11). But Tugendhat has on several occasions claimed that one can be nothing but a Humean in this question: "Nur Gefühle sind Handlungsmotive" (1997, p. 123).

<sup>6</sup> 'General term' and 'singular term' are the correct semantic concepts. Tugendhat uses 'predicate' instead of 'general term' for simplicity's sake, even though this is a syntactic concept (cf. Tugendhat 1976, p. 52n10). There have been discovered human languages without grammatical predicates, but not without the semantic general terms.



The structure of human cooperation is therefore radically different from that of other species. That brings us to the second idea that Tugendhat has imported from Aristotle in his account of universalism. There is a passage in the direct vicinity of the already mentioned description of our condition as *zoon logon echon*, where Aristotle explains that human communities are created through a common view on justice:

"it is a characteristic of man that he alone has any sense of good or evil, of just and unjust, and the like, and the association (*koinonía*) of living beings who have this sense makes a family and a state" (Politics 1253a15)

According to Tugendhat, the way Aristotle uses *koinonía* is similar to what he himself means when he speaks of "collective actions" (cf. Tugendhat 2007b:147). *Koinonía* is variably translated as society, fellowship, partnership or association, but what is at issue here is primarily the collective *taking part in something*. Perhaps "project" connotes better what both Aristotle and Tugendhat have in mind, since this concept involves an end that is perceived as good. A project is something that one has to actively promote in order to be a part of it. It is not unusual to see morality described as a "cooperative project", but then again; this makes it sound like morality is some deliberately designed and executed enterprise. Anyway, the central idea is that the cooperation of humans is based on an implicit or explicit acceptance of the ends that this coordination achieves (cf. Tugendhat 1979, p. 266).

It is this anthropological structure of collective actions that is supposed to reveal how universalism has an origin that can be described neutrally or in a pre-moral sense. It is, at first glance, not easy to see how he can extract this from the cited passage from Aristotle. The passage seems to say that only beings that already *have* a common concept of what is good and just can be joined together in a collective action. How can he use a quote that refers to people that have an understanding of the moral good when he is trying to explain this very concept? Our notion of moral goodness is supposed to be the result of the structure of collective actions, and now Aristotle tells us that there can be no collective action without a shared concept of the good.

However, Tugendhat's account can be fitted with the quote if we notice that he has equated Aristotle's talk of "perceiving something as good and just" with the *will* of the agent (cf. 2007b:147). So the precondition for *koinonía* is that there is agreement amongst the wills of the various participants. Charles Taylor seems to be talking about

the same phenomenon when he speaks of "common spaces", where actions are performed by agents that have reflective knowledge that everybody pursues the same goal:

"Their focus is common, as against merely convergent, because it is part of what is commonly understood that they are attending to the same common object, or purpose, together, as against each person just happening, on his or her own, to be concerned with the same thing" (Taylor 2007:187)

This can be illustrated with the example of some friends going to the cinema (avoiding the harshness of a chain gang this time). The "common space" or "collective action" is constituted through the fact that these friends did not just happen to meet at the entrance. Everybody perceived it as a good thing to hang out *together*, watching the same story and perhaps discussing it afterwards. This makes it a mini-society instead of just separate members of a movie audience. In a *koinonía* like this group the members have an *agreement* on what they are doing and why. This agreement is what keeps the group together. In order for some friends to agree to a trip to the cinema they all have to believe at some level that there is something good about this action, good enough to perform it. Only if all want to pursue the same goal can the collective action emerge.

It is this notion of agreement that Tugendhat wants to retain from Aristotle. The idea is agreement as a *precondition* for, and not a *result* of, the moral community. A notion of volitional agreement is involved in structure of collective actions as such; it is therefore also involved in the structure of the common practice of morality. This is not yet a full account of universalism. So far we have not been given an explanation of why the will of all must count equally. We have seen that the moral practice should be understood in terms of a "collective action" where several wills are joined, but we do not know why one way of joining wills is more justifiable than others.

### *3.3 Agreement or Force?*

In the previous section we saw that humans can decide whether to cooperate or not. What does this have to do with universalism? Tugendhat's thesis is that universalism denotes one of two extremes when it comes to structuring a collective action. In his terminology; a collective action can either be symmetric or asymmetric. Symmetry and asymmetry are meant to be normatively neutral and structural notions that replace

universalism and egoism. Why does he need this pre-moral understanding of ways of structuring actions?

This descriptive take on interaction avoids using prescriptive terms in order to understand further prescriptive terms. If Tugendhat depended on notions that were already part of the moral practice the understanding of universalism would amount to an axiomatic scheme. The moral sphere would be understood as governed by a special set of laws that are internal to this normative realm. Tugendhat's previous account of universalism suffered from this kind of problematic self-containment, since instead of a descriptive account of human action he offered a definition; "eine Moral begründen heißt, sie allen gegenüber gleichermaßen begründen" (1997a, p. 68.). The standard of universalism was explained as definitional of the moral practice. That may be so, but the question is why.

His final solution - that we are currently looking at - is an attempt to demonstrate that universalism can be understood without appealing to the concept of morality itself, which would be a circular explanation. Such an account would amount to saying that moral norms must be in the equal interest of all because that is the way morality is defined. In this final account it is rather to be explained in terms of the possibility of symmetric interaction. What does it mean that a collective action can be structured symmetrically or asymmetrically?

If the action tilts against asymmetry it means that the common end is decided by *one* of the participating agents, and the others have to adjust their will to his or her decision. This is what Tugendhat describes as a relation that rests on *force*: "Wenn die gemeinsame Handlung so aussieht, daß einer das Vermögen hat, das Wollen der anderen zu bestimmen und erzwingen, heißt das, daß er Macht über sie hat und daß die gemeinsame Handlung auf dieser Macht beruht" (Tugendhat 2007b:143). An asymmetric collective action, then, is an action with an end that only some of the agents actively want to promote. The others find that they have to promote the end only in order to avoid some worse consequence.

Let us again use the group of moviegoers to illustrate what Tugendhat has in mind. Imagine that Tracy has taken the responsibility to order movie seats for some friends, although they didn't decide on any particular film. She happens to know which film the others would most like to see, but it is also possible for her to pretend she didn't know and order tickets to her personal favourite instead. There is,

then, a project involving more than one will; yet only one option can be chosen. Tracy would put herself in an asymmetric relation if she ignored the will of the others, and let the decision be based on her will alone. The others would then have their own wills devalued.

Of course, the others could decide to go home instead. Tracy's power was limited to deciding the specific end; she could not decide whether the others must to participate at all. That would have been an additional case of asymmetry. There are two possible levels of asymmetry, then: One level concerns the influence one has when it comes to *participating* in the collective action at all. A second level regards how the various agents can influence the *end* of the action.

If the collective action is to be symmetric, however, Tracy cannot ignore the others: "Nicht einer entscheidet und bestimmt, sondern alle entscheiden gemeinsam, und das heißt alle *gleichermaßen* dazu beitragen, wie gehandelt wird" (Tugendhat 2007b:144). The notion of moral universalism has its root, then, in the extent to which the agents can influence the decision. At the asymmetric end of the scale one person decides. At the other end all people decide equally. Obviously, this resembles the idea of democracy - the thought a decision is legitimate only when it is the result of a process where the will of each counts equally. But the concept of symmetry is completely detached from any concrete social institution. It is conceptually prior to democracy, since it is meant to work as the model that would justify such a practice. Symmetric collective action is a possible structure of all kinds of cooperative undertakings; from arranging movie seats to making laws for a country.

I mentioned earlier that Tugendhat has arrived at this solution through use of both Aristotle and Habermas. We are now going to take a short look at how Habermas fits into this picture, since it sheds some further light on the notion of symmetric action. Habermas is only mentioned in passing in this context, but the reference is significant. In a reply to a critic Tugendhat remarks in parenthesis that he probably got the idea of seeing the matter as a basic anthropological alternative between force and agreement from Habermas (cf. 2006b, p. 309). If we knew what part of Habermas' theory Tugendhat is referring this could give his concept some helpful additional context.

Tugendhat is most likely referring to Habermas' concept of "communicative action": "[Ich] spreche von *kommunikativen* Handlungen, wenn die Handlungspläne

der beteiligten Akteuren nicht über egozentrische Erfolgskalküle, sondern über Akte der Verständigung koordiniert werden" (Habermas 1981:385). To be sure, what Habermas means by "Akte der Verständigung" involves a more comprehensive theory of linguistic and rational agency than the mere structural and minimal conception Tugendhat wants.<sup>7</sup> But the main idea is clear; our social interaction can be coordinated in such a way that the possibility of agreement regulates our pursuit of private ends.

The alternative of communicative action therefore involves introducing a qualification on all of our actions; that "Alter seine Handlungen an die von Ego anschließen kann" (Habermas 1983:145). This condition is only satisfied insofar as one does not take advantage of others by reducing them to vehicles to reach personal benefit. Doing this would be what Habermas calls "strategic action". This is the social alternative to communicative action, and does not place qualifications on how one reaches one's goals. Strategic action is successful if some end is achieved by calculating how others will act, it does not matter whether this involves disregarding the will of others or not.

If force is defined as the ability to make others do something other than what they perceive as good, then strategic action is compatible with force. Communicative action is not compatible, since its condition of satisfaction is agreement. So the alternative of communicative action is equivalent to the option of symmetry. This way of structuring actions is not only egalitarian, but is also irreconcilable with particularism. To limit one's disposition to reach agreement to a special group means not respecting the will of another group: "Bei einer Diskrimination von Außenstehenden (einem Sichnichteinlassen auf einen Begründungsdialog mit ihnen) würde Macht an die stelle von Einvernehmlichkeit treten, genauso wie bei einer Diskrimination nach innen" (2006b, p. 309). By being the opposite of force, then, symmetric actions must both maintain equality among members of the community and a respect for *all* possible cooperative beings that might be affected. If Tracy had to choose a film for her whole school class she could not respect only the will of those that belong to her closest circle of friends.

So in this section we have seen how Tugendhat identifies the content of universalism as a special configuration of wills in social action. But so far nothing

---

<sup>7</sup> Tugendhat has discussed Habermas' concept extensively in Tugendhat 1985, where he is particularly annoyed that Habermas is mixing a *structural* description of social action with an analytically independent *attitude* condition of consent-orientation.

normative about the concept has been demonstrated, it is simply one of two alternatives. The fact that there are two extremes on a scale does not tell us anything about why one of these extremes would be perceived as morally better. This was precisely Tugendhat's intent: It is supposed to illustrate a pre-moral option that only humans have. And yet it is an alternative that "unmittelbar zur Moral hinführt" (2007b:143). Where does the switch from descriptive to prescriptive occur? There are two alternatives, why is it so obvious that the egalitarian end of the scale is morally better than the other?

### *3.4 A Reciprocal Demand*

Both Kant and Habermas can explain this by recourse to reason. According to Kant we have to adhere to this standard in order to be autonomous. Habermas rejects Kant's transcendental reason, but still thinks that universalism is necessary to be autonomous, speaking of "Reinigung des Willens von heteronomen Bestimmungen" (2004, p. 340). A discourse of equal respect is the only way to get an outside perspective on oneself and rid oneself of prejudice. However, arguing in terms of freedom of the will in this way makes very limited sense: Only if one has already decided that all must count equally can one perceive a self-centred ego to be a problem. Tracy might experience that a dialogue with others makes it easier to pick a film without being determined by prejudices about what others think. But if that is the case, then evidently she *wants* to make an impartial decision. This will is an explicandum; it has to be more than just a whim. How does universalism become a normative notion by mere reference to our empirical interests?

We know already that Tugendhat does not think that morality is to be explained in terms of the autonomy of the *individual*, but rather the *collective* autonomy of the members of a community. The prescriptive nature of universalism is identified as a natural consequence of the way collective actions are structured. The symmetric alternative will be mutually demanded "sobald die Individuen ihre Situation so verstehen" (2007b:147). When agents understand their own situation as disadvantageous in relation to others a transformation takes place; the asymmetry is no longer neutral. If the participants in collective action recognise that the end is being decided against their will they no longer have any interest in participating. If their participation is necessary this creates a *pressure* to define the end in such a way that

all can agree to it. One alternative is now suddenly the precondition for organising a collective action at all, unless one of the agents has the means to force his or her will through.

The normative pressure towards equality arises from the fact that a collective action cannot be achieved without agents that want to pursue the same goal, that pull in the same direction. The basic structure of collective actions contains one end and many wills; the normativity of universalism is the force that arises out of the fact that a number of wills cannot be joined and directed towards a goal on unequal terms without coercion. When an agent recognises that his will does not count as much as any other he will demand a reason: "Why are my wishes of less significance than those of others?" If this cannot be answered in any satisfactory way he will not have a reason to take part in the collective action. He'll leave if he has a choice.

Let's go back to the example of going to the movies. If Tracy had arranged tickets to a film that she knew Richard had seen and disliked it would have been an asymmetric act since Tracy devalued his wishes. Richard might not get upset with her if he is unaware that she has done this, that is, if he thinks she did not know that he hates this film. Richard would not resent Tracy for what he believes to be an accident. He may of course get a little but frustrated, but resentment involves the judgement that someone has deliberately disrespected the will of the offended. But imagine he discovers that Tracy knew very well that he didn't want to see this film. It is plausible that once he knows that she does not respect his wishes he will find no reason to take part in other projects with her. This may have a further and devastating effect: If the others of the group see that Tracy does not treat people with respect they know that they might suffer next. This will make them shun interaction, as people do not want to be used as mere instruments.

This shows that the normative concept of universalism is not dependent on some metaphysical ideas or specifically Western intuitions about moral goodness. It can be made sense of within a general anthropological framework that starts off from preconditions of free will and unforced social interaction. Tugendhat presents the notion of symmetry as the real alternative to coercion, and we have seen how this makes it constitutive of the moral practice. Those that do not adhere to the standard of universalism rely on having the force to disrespect the will of others, they do not think

of their action in terms of what can be justified. This may work for some time, but they will eventually have to go to the cinema alone.



#### *4. Justice and Discrimination*

We are now going to look at how the standard of universalism relates to issues of justice. A just distribution must be defensible on terms of universalism, but justice differs from the idea of universalism in that it contains the notion of proportional distribution (cf. Tugendhat 2007b, p. 146). The task of this chapter is to look at the concrete ways in which Tugendhat perceives the standard of universalism to govern distribution of rights and goods.

Central to this chapter is Tugendhat's description of what he calls "primary" and "secondary discrimination". In 4.1 it will be argued that a closer look at the term "primary discrimination" reveals that Tugendhat is ambiguous as to whether one has to be human to have moral rights or not. His focus on cooperation leaves him with difficulties concerning those that cannot cooperate. In 4.2 it will be argued that Tugendhat's concepts of "primary" and "secondary discrimination" are misleading terms, considering that what he wants to describe are not to levels or stages of the same phenomenon.

##### *4.1 Cooperation and Universalism*

In the previous chapter we saw how the moral standard of universalism is understood as constitutive of any moral practice. But we know that there have been and still are many moral practices in which all are not granted equal rights. Discrimination of women, blacks, Jews and the lower castes are widespread occurrences. Tugendhat calls this "primary discrimination" (cf. 1993, p. 375). The common denominator in these cases is that some descriptive feature is taken to have normative relevance. People are of unequal worth from birth on because of their colour or religion, for instance. How can this be justified without a belief in a "higher truth" that transcends our mutual interests?

It can't, since referring to a natural inequality as the source of normative inequality is a "naturalistic fallacy" (cf. Tugendhat 1993, p. 375f.). We must be clear about what Tugendhat means by this. He is not thinking of a "naturalistic fallacy" in the way G.E. Moore formulated it in his *Principia Ethica*: "But if [one] confuses 'good' ... with any natural object whatever, then there is reason for calling that a naturalistic fallacy" (1903, § 12). This has been held as warning against interest theories up until today; "you cannot derive an 'ought' from an 'is'". Kantians like Karl-Otto Apel and Christine Korsgaard are especially fond of this meta-ethical law,

because it is seen as validating the project of grounding normativity in reason itself. Echoing Moore, Korsgaard writes: "The normativity of an ethical concept cannot be derived from any non-ethical concept, so no ethical concept can be completely analysed in terms of a natural or factual one" (1996, p. 66).

This is sometimes taken to be the same as Hume's Law, although he had a different point in mind.<sup>8</sup> Hume says that it is "altogether inconceivable, how this new relation can be a deduction from others, which are entirely different from it" (1888, p. 469). But using this as a reason to justify moral principles *a priori* amounts to taking Hume's advice the wrong way. Staying out of the empirical realm only seemingly avoids the problem by, since this strategy seeks only a different kind of "is". The problem remains. Hume's Law does actually not imply that we have to enter some other realm for the justification of moral principles; we should rather clarify what it takes for a statement to have normative impact. It was meant as a warning against moralising tendencies in ethical theory. He saw that it is fallacious to conjure up normative ideals while leaving the very nature of normativity unexplained. But this does not mean that normativity relates to something different than our empirical will, on the contrary.

Only when one adds some volitional identification with a moral law issued from some "higher authority" is it possible to understand unequal rights as the justifiable result of natural inequalities. If all believe in this authority the unequal rights that the authority prescribes are not forced upon them. This point is dependent on the understanding of universalism that was elucidated in the previous chapter. Since morality is the opposite of coercion it would be immoral to discriminate against people on the basis of a belief they do not identify with. So the rejection of all primary discrimination is basically just a rephrasing of what universalism means. However, it does provoke the question of where to draw the line. Can we reject equal moral rights to animals? Isn't being a human being also a descriptive feature?

Tugendhat admits that we need *one* descriptive feature in order to delimit the moral sphere, but being human is not the relevant attribute. He speaks of morality as a community of "cooperative beings" (cf. 1993, p. 193). So in order to be the subject of

---

<sup>8</sup> Apel speaks of Hume's Law when he is actually more in tune with Moore. In an attempt to show why transcendental-pragmatics does not violate the "is-ought distinction" he says: "Die Akzeptierung der moralischen Grundnorm der kritischen Kommunikationsgemeinschaft hat, sofern sie notwendigerweise vorausgesetzt werden muß, nicht den Charakter eines Humeschen 'Faktums', sondern den Charakter des Kantischen 'Faktums der Vernunft'" (1973, p. 417). Hume would not be convinced.

equal moral rights and duties one has to belong to this cooperative community. One does not really have to be able to cooperate with others; the disabled also belong to this community, according to Tugendhat. This means that humans that are infants, handicapped or temporarily indisposed still belong to the community of cooperative beings. Now why does he choose the feature of cooperative being? This has to do with his understanding of morality in terms of a "collective action", which depended on the cooperation of all. Animals cannot choose to cooperate, since they just do what their instincts say.<sup>9</sup> In order for there to be a reciprocal moral "must" involved there has to be a choice involved. The concept of obligation neither concerns laws that we are fully determined by nor maxims that we are unable to adhere to. But what about humans that cannot cooperate; how can our moral duties towards these be explained?

Tugendhat simply states that "alle, die zur Kooperationsgemeinschaft gehören, aber selbst nicht oder kaum kooperieren können, lediglich um so hilfsbedürftiger sind" (1993, p. 377). This seems intuitively fair, but how can one say that a person belongs to a community that is united in virtue of an ability that this person lacks? It is difficult to accept Tugendhat's way of presenting the solution, since it is not sufficiently in tune with what it means to be a member of a cooperative community. Merely stating that humans that cannot cooperate are still part of the community of cooperative beings seems to rely on the descriptive feature of belonging to a species, but that was to be avoided.

In accordance with his understanding of morality as a cooperative enterprise he needs to understand the moral rights of those that cannot cooperate by a detour that may clash with our intuitions: Only insofar as those that can cooperate demand of each other that all humans are to be normatively equal are the rights of the disabled secured. The reason saying this seems awkward is of course that it makes the rights of those that cannot cooperate dependent on the will of those that can. Moral obligation towards infants will be contingent upon whether the cooperative beings of the community happen to care about infants.

Most likely they will care, but the point is that the principle of universalism does not stretch directly to our relation to the infants or severely handicapped. It is possible to envision a community that was morally legitimate according to universalism and yet did not have equal respect and concern for those that were

---

<sup>9</sup> For Tugendhat's suggestion as to how animal rights can be incorporated into his concept of philosophy, see 2006, p. 29.

unable to cooperate. But this is the only way to explain their moral rights on terms of symmetric contractarianism. Nobody has intrinsic moral rights from birth; they are constituted through reciprocal wants. This is of course incompatible with the formulation in the Declaration of Human Rights that speaks of "the inherent dignity and of the equal and inalienable rights of all members of the human family". The way Tugendhat has explained the matter rights are not *inherent* but *received* from the community of cooperative beings. This cannot be solved by simply smuggling those that cannot cooperate into the class of cooperative beings, like Tugendhat seems to have done.

#### *4.2 Equal Respect and Unequal Distribution*

As the quote above stated, those that cannot cooperate need more help than those that can. This brings us to Tugendhat's notion of "secondary discrimination" (cf. 1993, 378ff.). This term refers to rules of distributive justice, where unequal distribution is needed in order to respect the moral equality of all. The concept of 'justice' is different from 'universalism' in that it incorporates the idea of *proportional* equality, which is unequal treatment that is in accordance with the normative equality of all that universalism prescribes. The reasons that speak in favour of giving some more than others can be several: They may not *need* more, like those that cannot cooperate, but have perhaps worked *harder* and should be compensated, for instance. What is important when it comes to secondary discrimination is to see how it relates to universalism. Inequality in distribution is a means to respect the moral equality of all. This point clarifies how universalism is incorporated into the distributive mechanisms in society. Proportional inequality does not contradict the equality of universalism but is a necessary consequence.

But there is a difficulty with this notion of secondary discrimination. It is made out to be a different kind than primary discrimination, when in it should rather be seen as derivative aspect of the same matter: Illegitimate secondary discrimination can always be revealed as primary discrimination. What Tugendhat describes as illegitimate secondary discrimination is only a means to operationalise primary discrimination. This can be illustrated in the following way: Let's say one group decided to grab twice as much of the community's goods as another group. If they could force the others to accept this distribution then such an action would not need to

be justified. But this is not the relevant case when looking at secondary discrimination as a contrast to primary discrimination: We are trying to see what kind of inequalities that can be justified. The point in looking primary discrimination on modern premises is to show that all inequality is arbitrary, that is, discrimination when it comes to moral rights cannot be justified. Similarly, the point in looking at secondary discrimination is to see what reasons could be used to promote it. So we should consider a case where the one group can't force others and need moral reasons. In this case they would have to appeal to the rules of proportional justice. And these rules could either be set by a "higher authority" or based on universalism. This means that the level of primary discrimination decides the result.

My argument is that the two kinds of differentiation referred to as "primary" and "secondary" are in reality connected in a way that is only obfuscated by Tugendhat's way of presenting the matter. The relation is simply like this: Illegitimate secondary discrimination is a way of putting primary discrimination into effect. Since proportional distribution can be an *instrument* to carry out ideas of moral inequality it is misleading to describe them in terms of "primary" and "secondary". It appears that the only reason Tugendhat compares primary to secondary discrimination is that they both contain a notion unequal distribution ("Ungleichverteilung"). It would be more comprehensible to simply speak of a configuration of moral rights and proportional justice. The first notion refers to a *standard* and the second to the *means* to maintain this standard in practice. The fact that both notions can involve a differentiation or discrimination is not the essential thing, since we are not speaking of a differentiation of the same sort of things in the two instances. In one we are speaking of moral rights and in the other of the distribution according these rights.

However, once this is cleared up we can appreciate the importance of having a proper justification of universalism itself when discussing matters of justice. Simply saying that justice refers to a balance between individuals is empty, since one could argue that all distributions are balanced according to *some* standard. The question in the previous chapter was why this standard has to be universalism. The thorough elucidation of this standard allows provides an anchor for our ideas of just distributions, relating unequal distribution to an idea of equal moral rights. It makes it comprehensible why those who have offered their sweat should have more, and why those that already live on the sunny side should receive less. The different

mechanisms of justice *even out* inequalities and push the collective action towards the symmetric end of the scale.

But symmetric action and egalitarian justice do not stand and fall together. That is; egalitarian justice is not necessarily the resulting outcome of a situation where all could decide equally how the distribution of goods should be. That is; universalism understood as symmetric collective action and the egalitarian notion of justice are not on the same level. What is meant is that a cooperative arrangement can be justified according universalism, while not satisfying egalitarian concept of justice. Restoring balance between individuals is not the sole function of universalism: "Gerechtigkeit ist nur ein Aspekt des Guten" (Tugendhat 1993, p. 385).

Tugendhat clarifies this conceptual differentiation by contrasting it to Rawls' "Justice as Fairness" which involves a "difference principle". According to this concept, the better situated may receive more if and only if it improves the situation of those that are least advantaged. Tugendhat points out that this is an aspect of utilitarianism and not justice; it does not create balance but imbalance in order to improve the greatest happiness (cf. *ibid.*, p. 386). This means that it is inaccurate of Rawls to identify the difference principle as a means to maintain the egalitarian concept of justice. The fact that all could identify and accept the difference principle makes it morally *legitimate*, but it does not thereby make it into an aspect of egalitarian *justice*.

Rawls could perhaps defend himself by claiming that he has *defined* justice by his two principles of justice. But Tugendhat is right in saying that this would collapse different normative notions in an unfortunate way. It may sometimes be necessary to weigh aspects of justice against matters of utility, creating a result that is morally good without satisfying an egalitarian rule of distribution. Without Tugendhat's thorough elucidation of the foundation of universalism itself - that was discussed in the previous chapter - it would be difficult to comprehend how the different normative notions of egalitarian justice and utility could be equally legitimate. We have seen that non-coercive agreement among wills that is the final source of legitimacy, and this does not predetermine what kinds of distributions would be legitimate.

## 5. Moral Justification

In the three preceding chapters we have been looking at three different questions: What is Tugendhat's concept of morality? What is the origin of the standard of universalism? How is universalism to be applied in matters of distribution rights and goods? In this final chapter of Part 1 we are going to see how the practice of morality can be justified to people that do not understand why one should participate in such a practice. That is; we know why Tugendhat thinks we must understand morality in terms of universalism, and now the issue is how one can explain the agents' volitional relationship to such a practice.

By developing a general concept of moral obligation Tugendhat has provided a clarification of *what* needs to be justified at all. Morality is made up of mutual imperatives backed by social pressure, and it is this pressure that needs to be justified: "Was begründet wird, ist gerade, diese Normen zu fordern und d.h. ihnen gegenüber moralischen Gefühle zu haben" (Tugendhat 2006a, p. 21). So what we must be prepared to justify are the reactive attitudes that we have connected to certain norms. This makes it clear how one can speak of the moral emotions in moral discourse without making it seem like our moral language is just "will against will".

If one pursued this latter idea one would be in agreement with Charles Stevenson, who said that two people involved in a discussion about norms are merely having a disagreement in interest (cf. 1937, p. 27). He claimed that norms cannot be justified because imperatives revealed preferences and not beliefs. Hence to present someone with a moral imperative is equivalent to saying "*I approve of this; do so as well*" (Stevenson 1944, p. 21). In such an account the moral language is only capable of revealing the fact that we have a certain attitude, and by conveying this we can aim to influence the will of the other.

However, the moral attitudes and the idea of objective justifications stand in a complementary relationship, according to Tugendhat: It is our emotional disapproval that creates a need for justification. The question that has to be answered when it comes to justification of a norm is why one should have the reactive attitudes identified in the Strawsonian Triad to this kind of behaviour. This way of relating the process of moral justification to affective sanctions highlights the fact that we are dealing with a question concerning "what to do" not "what is true". It is the attempt to provide a *practical justification*, and this must be understood in light of a theory of

action: "Ich glaube, daß wenn wir von der Begründung einer Norm sprechen, das dem näher kommt, wenn wir von der Begründung einer Handlung reden" (Tugendhat 1997b, p. 16).

The point is that a moral norm cannot be justified analogously to the way we justify empirical statements. A norm is only justified insofar as people find a reason to limit their freedom according to it. This is not the way an assertion is justified: "The earth revolves around the sun" can be supported by empirical facts, and can be demonstrated to be true no matter whether people like it or not. But there is no evidence in this sense that can make us identify emotionally with a moral norm. That does not mean that the norm cannot be justified at all; there can be good reasons in terms of *motives*. By why do we have to our reactive attitudes to each other in the first place?

### 5.1 Why Justify?

In order to illustrate why one needs to justify norms at all Tugendhat sometimes speaks of the child that wonders why the parents get upset when it has violated some norm; "warum werdet ihr so böse, wenn ich so handle, warum erregt ihr euch so?" (1998, p. 94). This is an important heuristic. First of all it brings the central problem to the fore; how can we justify our emotional disapprobation? The child expects a reason that shows why this reaction is appropriate; it does not want to be subject to random negative emotions. A justification of the reaction can show how the norm is something that the child should *identify* with, and not just some attempt to bend its will in the interest of some end that it does not recognise itself.

But we must be wary of this child example, though. It has an unfortunate potential to be misleading, since the situation of justification to a child is different from the relations between ordinary adults in an important regard. Using the relation between the child and its parents is too similar to the relation of the citizen and the state. Both the child and the citizen live within a social arrangement that they cannot choose freely. But an important feature of morality is precisely that we can choose; once we consider ourselves as outsiders to the moral community we are more or less immune to the sanctions. A person that believes that morality is symptomatic of slave mentality is not subjected to the moral norms. No sense of guilt is ignited in this agent



when others react with resentment, because he or she does not accept the underlying principles of the norm.

This point is often blurred out in Tugendhat's writings, even when it is not connected to the example with the child. As for instance when he says; "in einer Moral befindet man sich, ob man will oder nicht" (1999a, p. 163). This makes it seem like moral norms limit our freedom no matter what we think of the standards of our community. But this is a dubious approximation to political thinking, giving the impression that we live inside a morality in the same way we live in a country regulated by law. An earlier quote makes this false assimilation explicit: "Den Sitten und Rechtsnormen ihres sozialen Systems hingegen sind die Individuen unterworfen, ob sie wollen oder nicht" (1981a, p. 75).

In these quotes it is not sufficiently understood how the situation of moral justification is exactly opposite from the sphere of law: A law has to be justified *because* it limits our freedom whether we like it or not, but a moral norm cannot limit our freedom *unless* it considered justified. Norms must be justified in order to attract volitional identification in society, not in order to explain some *de facto* limitation in freedom. That is; we do not need a justification as to why our freedom *is* restricted, but a justification as to why we should *want* to restrict our freedom. At some level we must willingly bind ourselves to the moral norms, accepting the limitation in our freedom of action that this represents. It is this act of restricting our own freedom that needs justification. This is a critique that is immanent to Tugendhat's writings. Remember the quote where he claims that the sanctions are dependent on a moral self-conception: "Sie kann nur diejenigen treffen, die sich als zugehörig zur moralischen Gemeinschaft verstehen" (1989b, p. 322).

Thinking of moral and legal sanctions in terms of a mere gradual difference in degree of institutionalisation does not capture the distinction between internal sanctions of conscience and external sanctions of law; they work according to completely different mechanisms. Think of the chain gang example that was used to illustrate moral self-restraint; each prisoner holds a key, and the reason he doesn't unlock is that he wants all to remain within certain bounds. But real prisoners do not have keys to their shackles, and this makes their loss of freedom a very different matter.

Wavering on this point is dangerous, because if the moral norms are said to restrict our freedom "whether we like it or not" we do not yet know why the project of moral justification is necessary. It fuels the false idea that justifying moral sanctions is *moral* demand: "Es ist die wechselseitige Empörung (der soziale Druck) die sich als illegitim erwiesen kann" (Tugendhat 2006b, p. 277). The idea is that moral norms must be justified because it is immoral to restrict others without reason. But it is circular to explain the need for moral justification by reference to a moral demand. We are taken into a regress if we understand the process of justification of moral imperatives in light of one "super imperative", like "Do not restrict others freedom without justification".

The only way to avoid this is to reject the claim that the necessity to justify our sanctions is a moral one. It is a pragmatic demand, since otherwise nobody would care about the resentment they are faced with. Against Tugendhat it must be maintained that the justification of moral norms is not meant to show that our reactive attitudes are morally legitimate in this situation, but is rather a precondition for the possibility of moral restrictions at all. If people were not presented with good grounds to identify with some moral norm they would not experience any obligation to it.

### *5.2 Why Accept a Norm?*

The next question is what kind of reasons one could use in order to make others accept the moral norm. The justification could in principle appeal to anything, as long as it gives the agent a reason to comply. If we understand ourselves as children of the Good Lord we have a reason to adhere to his commands and a reason to resent those that transgress them. A justification could then amount to appealing to the relevant passage in the Bible, showing that this is the real will of God. But the problem concerning the modern concept of morality is that we cannot presuppose any particular belief, so the reciprocal reasons we give each other must relate to our ordinary interests.

So norms must be justified by recourse to what the members of society actually want from each other. That is; the interests provide the content to the norms. If one does not believe in some "higher authority" there are no grounds to believe that the moral norms are "out there" somehow, independent of our actual needs and desires. As Stemmer writes for provocative effect: "Wenn niemand daran interessiert

wäre, nicht getötet zu werden, dann wäre dass Tötungsverbot nicht Teil der Moral" (2000, p. 193). This may stand in tension with our pre-reflective moral consciousness. But it is likely that norms have a more tangible appearance in everyday perception than can be made philosophical sense of. But once the matter is reflected upon it should not cause too much distress. It is not discernible how some physical, metaphysical or psychological "reality" status should matter to us anyhow. What matters to us is whether the normative system is considered justified in a way that relates directly to our own will. Is it justified *to us*?

What does a justification by recourse to interest imply? Tugendhat is usually satisfied with using the concept of interests in a very loose fashion. It has the function of being the terrestrial alternative to "higher truths". But are all interests morally legitimate? Formally no, substantially yes. That is, the interests must have a *reciprocal configuration*, which is the formal criterion, but the substantial content of these interests is left open. Why must the actual content be left to the agents themselves? Tugendhat refuses to claim any privileged standpoint, so a list of morally legitimate interests must result from an actual intersubjective process of justification (cf. Tugendhat 1997a, p. 39). This means that the concrete rights and duties cannot be static or deduced once and for all. The process is always ongoing and the outcome changing with our revised interests. In this sense, all interests are potentially legitimate.

There can of course be different ways that a norm is in the interest of the agents. A norm may be in the interest of each member of society, but to a greater advantage to some. An underpaid employee still gets paid, and both contractor and worker make a profit. According to strict contractarianism this relation is moral as long as it is to mutual advantage. Gauthier and Stemmer think that moral justification only amounts to showing how it is in our self-serving interests to be moral. This view of the relation between interests and norms does not allow the standard of equal respect for all to decide whether a norm is the legitimate or not, since "the object of rational co-operative choice must be an optimal outcome" (Gauthier 1986, p. 117).

This has the consequence that those who cannot contribute to an "optimal outcome" are left with no moral protection. As long as it is not in our self-interest to cooperate we are under no obligation to do so. Stemmer only speaks of relations of strength, where we are obligated solely towards those that could potentially hurt us:

"Wenn A nicht in der Lage ist, B zu verletzen, dann hat B keinen Grund, sich vor der Verletzung durch A dadurch zu schützen, daß er eine Verpflichtung übernimmt, A nicht zu verletzen" (Stemmer 2000, p. 210). Those of equal strength have prudential reasons to take on an obligation towards each other, but there is no reason to care for the weak. Now, the content of the norms that the strong adhere to may overlap with much that is prescribed by universalism (eg. don't lie or steal), but the obligation to these norms internal to the circle of people that find it beneficial to cooperate with each other. This is because equal respect within the community is a precondition for a rational agreement on norms, but it is only rational to make an agreement with those that matter to one's own self-interest.

The difference between universalism as presented in Tugendhat's concept of morality and the contractarianism of Gauthier and Stemmer is not immediately discernible. Tugendhat also understands morality as a cooperative project, and as we have seen his theory cannot make our obligation towards those that cannot cooperate comprehensible without a detour by those that can (cf. 4.1). Even though principle of universalism as interpreted by Tugendhat does not allow the use of force against other cooperative members, the principle does not apply to actions towards the disabled without further premises that concern the will of the able ones: It must be in the reciprocal will of the cooperative members to take care of those that cannot help themselves.

But the criterion "ability to cooperate" is quite different from "ability to hurt", which Stemmer operates with. My ability to cooperate can be very limited but still enough to grant me moral rights by appeal to the principle of universalism. The ability to hurt, on the other hand, presupposes far greater physical and mental skills. Now, the criterion of being able to hurt others is very understandable; we want to stop potential threats to our safety. If moral norms were only meant to secure this it would be an easy task to justify them to the agents that are restricted by them. But why should we accept norms that go beyond this principle of safety? Why should we accept a principle that demands equal moral respect for all cooperative beings, no matter their potential to hurt us?

### 5.3 Justification of a Universalist Concept of Morality

Tugendhat's conception of morality implies that legitimate norms are in the *equal interest of all*, not just to the advantage of each member. What we must appeal to when justifying the moral practice is not the maximisation of our interests but the possibility of not using others as mere means to an end. But how can an individual be motivated to enter a practice in which it must let the interests of all others count just as much as its own when it comes to justifying norms?<sup>10</sup>

Those that do not have the powers to fend for themselves have an obvious motive to enter such a practice, since it provides better protection for their interests. The question is not how one could convince the forlorn, then: "Die einzig interessante Frage ist, warum eine Person, die für die Macht optieren könnte, eventuell das Symmetrische vorzieht" (Tugendhat 2007b, p. 155). Importantly, the task is not to show that the strong and powerful *must* choose morality. As has been mentioned, he does not think that the philosopher has to demonstrate how morality can be construed as unconditionally necessary. His task is rather to show what is at stake when deciding whether to be moral or not. What does it mean to choose morality? And can regular moral behaviour of the powerful be made sense of?

I will discuss the answer Tugendhat provides in his latest essay "Der Ursprung der Gleichheit in Recht und Moral" (2007b, p. 155). Tugendhat identifies two possible motives, only one of which is a moral motive according to his definition of morality. The first task is to learn what Tugendhat means by each of the alternatives. The next question is whether or not this way of stating the alternatives implies a shift from his earlier view of the relation between the prudential interests and moral goodness. His previous accounts seem to say that the two alternatives should rather be viewed as different *stages* of motivation, instead of being contrasted as the moral and the non-moral motivation.

But let us start with what Tugendhat describes as the non-moral motivation. This is explained as when "symmetrisch zu existieren der Person befriedigender erscheint als die Machtausübung" (ibid.). This option entails that the person sees interaction on terms prescribed by universalism as "Teil des guten Lebens" (ibid.). This would be the Greek justification of morality, where our true prudential interest

---

<sup>10</sup> Given that Tugendhat's basic concept of morality already implies universalism we can disregard his differentiation between being moral in a formal sense and being moral in the sense of respecting the equal rights of all (cf. 1993, p. 90f.).

(to agathon) was said to consist in the pursuit of the moral goodness (to kalon). Why doesn't Tugendhat consider this kind of motivation to be of a moral kind?

The reason is that the moral practice is there to ensure that people live by the standard of universalism *even though* they might perceive their happiness to lie elsewhere. Tugendhat is in agreement with Hume that the object of morality is precisely to *compensate* for our selfishness and limited benevolence (cf. Tugendhat 1996, p. 325f.). The point is that morality is about what we are obligated to do no matter how what our preferences are. So those that are moral simply because it is more rewarding have a non-moral motivation since it is independent of the moral sentiments of guilt and resentment. Had our will to live in moral relations with others been strong we would not need to speak of an obligation, meaning we would not need the kind of sanction-based morality that he has been depicting. The ones that discover happiness in universalism have practical maxims with a moral *content* based on a non-moral motivation.

The specifically moral motivation is when the agent "die soziale Ächtung fürchtet" (Tugendhat 2007b, p. 155). People want to be recognised as good, and that implies respecting the reciprocal wants in the community. But this is an external restraint; morality is not something one genuinely wants, rather an impediment to take into account in order to stay socially afloat. So the motivation is social recognition, it is not directed at moral goodness itself. One is not moral because one has identified with the practice. But this implies that one has no reason to be moral when it comes to actions that can be performed in secret. Morality could then in principle consist of "freeriders", people that give the impression to have a moral character but in reality fake the moral sentiments (cf. Tugendhat 1997a, 43f.).

That morality could consist of mere freeriders may be perceived as a fault in Tugendhat's concept of morality. Should one not explain the fact that some people actually behave morally when unseen as well? It may be the case that our actual moral practice consists of a great many freeriders, but we seem to be short of an understanding of moral motivation if we cannot make sense of those that have a conscience. Tugendhat is fully aware of this: He claims that our conscience constitutes itself as we internalise the intersubjective demands. This makes us no longer fear just the actual rejection of others but also lack of worth in our self-

perception (cf. 2006a, p. 21). So our conscience is derived from the contempt expressed by others.

Moral shame is the transgressor's correlate to the contempt he faces. The feeling of shame is the feeling of being *devalued*. This is how Tugendhat defines the moral shame that arises in the transgressor: "Gefühl des Selbstverlustes in den Augen der (möglichen) Anderen" (Tugendhat 1993, p. 57). He claims that this definition is found in *Pride, Shame, and Guilt* by Gabriel Taylor (1985). There is a minor discrepancy between their accounts, however, seeing as she does not actually claim that the audience has to have an adverse judgement of the agent. One can feel shame even though one is seen with approval because one is degrading oneself by appealing to an audience on a lower level (cf. 1985, p. 64). Taylor says the notion of an audience is important mainly because it is a matter of getting an *outside perspective*; one is torn out of a pre-reflexive state and aware that one is in a state one should not be seen (cf. 1985, p. 66). In order to say that the agent loses worth in the eyes of the audience, as Tugendhat does, one needs the proviso that the agent cares about the opinion of the audience. This is implied when he speaks of "Applaus der Urteilsfähigen", that is, we want to be seen with approval by those whose judgement we trust (cf. 1997a, p.124).

The important aspect is that we feel that others perceive us as of no worth, or that they would if they had seen what had happened. Saying that our conscience is derived in this way from the actual moral norms in society may give the impression that we are determined by already given standards. There seems to be little room for autonomy in relation to the present moral imperatives. But Tugendhat does not think that people in general are satisfied with being recognised by their society. According to Tugendhat, we can assume with Adam Smith that people do not only want to be recognised but also *worthy* of recognition (cf. 1997a, p. 123).

A community may have norms that are considered good only because some inequalities have been redescribed in order to give the *appearance* of legitimacy. If one has reached the stage of wanting to be worthy of recognition one may try to convince the rest of the community that the common norms are in fact unjustifiable. This person would have to change the way people think and talk about those that are treated without respect: Discriminated groups are usually described in sub-human ways, making unequal treatment seem like proportional equality. A community never perceives of its own practice as immoral. The one that goes against this kind of

communally accepted discrimination is not governed by what is actually praised in society, but what he or she thinks is worthy of praise. This is an allusion to the Socratic idea of living in accordance with what one perceives as good and not just what is considered good. But it cannot be a social obligation to have this kind of autonomous conscience, since the social norms are just what one is supposed to have an autonomous relation to.<sup>11</sup> The person who silences his autonomous conscience in order to secure social recognition is moral by the standards of his community (but a coward as well).

But those that dare speak against communally accepted injustice are in the business of clearing up the matters of fact instead of giving people motives. This means that we are speaking of reasons in a different sense: "Das sind dann nicht mehr Gründe im Sinn von Motiven, sondern es sind Gründe für Aussagen" (Tugendhat 1997a, p. 48). The attempt is not to explain why one should be moral, but rather what is truly morally justifiable. In a religious community this could amount to explaining what the scripture actually demands, while in a modern moral community it would mean demonstrating what it requires to treat people as equals. Only here can we speak of the truth of moral judgements. There is a fact of the matter whether people are treated with equal respect or not.

So from this one can distil three different levels of moral motivation that could explain a preference of morality to force. The first was the fear of social rejection, which is compatible with being a freerider. The second implies having a conscience, which means feeling shame even when nobody is looking. The third level involves disconnecting shame from what is merely socially expected, and adjusting one's behaviour only in relation to behaviour that one personally thinks worthy of contempt. Does all this give us an answer to the initial question of why the person that could choose force would rather prefer moral relations?

It is actually the account of the non-moral motivation that is most convincing. Tugendhat's account of the moral motivation that I have just looked at relies on an agent that already fears social contempt. But this is avoiding the problem, since we were primarily interested in the person that "für die Macht optieren könnte". If he fears social rejection then this amounts to saying that he does not really have the strength to get by without moral relations. And by taking the next step and giving this

---

<sup>11</sup> For the relation between our intellectual honesty and intersubjective demands see Tugendhat 2007d.



person a conscience we were fading out the option of using strength instead of moral interaction.

The interest in the question was surely to see what kind of reasons one could have to *enter* a practice that implies putting a condition of equal respect on one's practical maxims. Note that this is specification of Tugendhat's initial question. As has been shown, he differentiates between being a member of the moral community and actually wanting to behave morally (cf. 2.1). Our discussion of the moral motivation has until now been concerned with the latter, since Tugendhat immediately brought in fear of sanctions. One can only fear sanctions if one understands oneself as a possible subject of these sanctions. I will now concentrate on why someone that could choose strength would want to join a sanctioned practice based on equal rights in the first place, meaning putting oneself under social pressure.

Evidently Tugendhat cannot demonstrate this on the basis of self-interested rational choice, since that would only get him the limited contractarian society that does not have room for universalism. What can he appeal to then? He has to return to his first alternative, the non-moral reason that was described as taking relations of equal respect to be something that has worth independently of sanctions. It implies an appeal to a certain *self-conception*: "man möchte nicht Tyrann sein und möchte sich deswegen auch nicht an einem System beteiligen, in dem die eigene Gruppe in einer tyrannischen Beziehung zu anderen stünde" (Tugendhat 1998, p. 108). Not being an oppressor is seen as the more fulfilling way of living, opening up for relations with people that would otherwise not be possible.

Choosing morality implies choosing a whole range of ways of interaction, several of which are so deeply embedded in our practical framework that to most people it could only be a hypothetical question whether one would like to continue this behaviour or not. This becomes evident when Tugendhat lists what is lost when one rejects the moral practice the way he understands it: "Es gäbe jetzt keine moralischen Gefühle, keine Disposition zu wechselseitigen Lob und Tadel und folglich keine internalisierendes Gewissen und somit kein Gefühl des Verpflichtetsein" (Tugendhat 2006a, p. 24). Rejecting the standard of universalism implies that others can be used as mere means, which would amount to reducing our social sphere to a sphere of tools. People would be more or less useful objects;

sometimes we would be used as such ourselves. Not wanting to be a tyrant, then, means wanting to live a life based on equal respect for all.

It seems more than plausible that someone who could remain uninjured in a Hobbesian state of nature would still prefer to have relations where we could interact in our basic human ways. The principle of mutual respect is necessary in all stable relationships, both our emotionally significant ones and to more distant acquaintances. The moral standard of universalism is a natural ingredient in a rewarding social life. But should we claim that this prudential motivation to be moral is an alternative to those that act morally out of a need for recognition? The need for recognition is surely the *result* of a will to live in balanced relations with others. So isn't it misguided to present our stand on "who we want to be" and fear of sanctions as alternative motivations?

Yes, since being a member of the moral community implies having made a decision on how to live a good life - at least implicitly. It is therefore somewhat unclear why Tugendhat has made the two motivational aspects out to be alternatives in this last essay. The non-moral motivation is conceptually prior to the moral motivation that is connected to social recognition. This fact has been a repeating theme in Tugendhat's writings, claiming that we must overcome the opposition between an ethical theory of happiness and a moral theory of principles (see esp. Tugendhat 1980). His solution as to how we can understand the moral principle universalism to be joined to our sense of the good life is summed up in this way; "wenn jemand moralisch sein will (oder dies in einem bestimmten ausmaß sein will), definiert er das Moralischsein als einen Teil seines prudentielles Wohls" (2003, p. 73).

It is of central importance to acknowledge this aspect of moral norms having to inhabit a place within our idea of the good life. If they didn't we could not understand the motivation of the powerful to limit their freedom according to the standard of universalism. So it is confusing when Tugendhat presents fear of social rejection and making universalism part of one's idea of the good life as alternative motivational sources (as he did in 2007b, p. 155). Instead it seems like the question of why some powerful person would prefer universalism should be answered by referral to *both* alternatives at once: He has as self-conception that implies not wanting to be an oppressor, and would *therefore* feel shame when he transgressed a norm that is prescribed by the standard of universalism.

The point from Hume about morality being a compensation for our limited benevolence must be understood in light of this way of arranging the alternative motivational factors. The need for recognition *intensifies* our will to make our behaviour conform to what we perceive as the proper way to live. Having decided that moral relations are worth pursuing is not enough to ensure that other non-moral ends do not trump this decision. The only way to ensure the necessary reliability is to react with affective sanctions when others betray their commitment to the moral norms.

#### 5.4 Summary of Moral Justification

In the previous section it was mentioned that one could understand the reasons of moral justification either as *motives* or as *evidence* for a statement. These two different understandings of "reason" point to the fact that the intersubjective justification of a moral practice is a two-tiered affair: On the first level we explain what is lost if one decides to relinquish one's moral existence. This is the justification of the moral "must", making it plain why moral norms should govern our behaviour at all. The second level involves clearing our prejudices or ignorance by demonstrating how some norm would contribute to normative equality.

If one wants to save a place for "ought" in morality it has to find its home on the second level (cf. Tugendhat 1993, p. 63f.). This is the level on which advice can be offered on how the moral norms should be *altered* in order to mirror the moral standard that is considered good: "Given that you want to adhere to universalism you *ought* not discriminate against blacks". This can be described as "epistemic" since it increases the knowledge of what would actually be in the equal interest of all. As we shall see in the next chapter, Habermas wants to restrict moral justification to this epistemic sort, and claims that obligation is a consequence of knowledge of this kind of reasons.

But it is the first kind of justification - we can call it "volitional" - that is of greatest concern to Tugendhat. Why enter a moral practice that involves treating others as ends and not mere means? This question was identified as involving our basic orientation towards others. It was in the end made into a question of "who one wants to be". To help further illuminate why it is appropriate to call this kind of justification "volitional" we should take a short look at Tugendhat's four levels of freedom of the will, where it becomes clear how morality can be a determination of

our self (cf. Tugendhat 2007c, p. 61ff.): The first level is self-control, our ability to put current desires on hold in order to reach some greater good. The second level is our ability to prioritise, to organise our wishes according to a self-conception. Morality is on the third level, and it represents one possible self-conception. The shared reasons of the community become a part of the prudential good, we understand ourselves as moral. Tugendhat finally mentions a fourth level, which is the autonomy that can be achieved in relation to the second and third level. We either shape our prudential or moral values in light of society's moral norms, or we can ask: What is *really* the best way to live and to treat each other? Tugendhat sometimes calls this "die Tiefendimension des Fragens", and it has been a central issue in some of his later essays (see esp. 1999b).

The philosophical project that we are involved in at the moment is placed on this fourth level, since it is asked for an account of the moral standards themselves. And the investigation that was conducted on this level leaves us with a clear perception of the relation between the second and the third level: Our moral self-conception (third) is dependent on an explicit or implicit decision on how to lead a good life (second). Why would one want to live in relations based on universalism if one could live a life based on pure self-interest instead?

The choice to not enter a moral community was shown to involve the loss of a great many features of our social lives that are generally highly valued. But this is no absolute justification; it is relative to our will to lead a certain kind of life. The option of entering the moral practice must be chosen if one wants to preserve the features of social interaction that depend on universalism; features like the reactive sentiments of guilt and resentment (cf. 2.1), stable relationships (cf. 3.4), justice (cf. 4), and the ability to back our negative reactions with reasons (cf. 5.1). In the preceding chapters we have seen how these different aspects of morality fit together, and when it comes to justifying our moral practice one must demonstrate that they fall together as well.

In Part 2 we are to find out whether Habermas and Stemmer are right to accuse Tugendhat of misidentified the basis of obligation. Has Tugendhat interpreted the moral "must" on wrong premises? Before moving on to the next part it is important to emphasise the way Tugendhat has justified our obligation as relative to our own decision. This shows that our obligation to be moral is not due to some pre-existing fact about us; it is not some feature of all humans like reason or our nature that

grounds our obligation. It is an obligation based on reasons that are sufficient for most people, but not for all possible collaborators. This represents a rupture with a way of seeing norms as either absolutely justified or not justifiable at all. This way of understanding the matter is captured in a quote from Wolfgang Kuhlmann (who favours the strategy of absolute justification):

"Das Argument ist einfach: Wenn Letzbegründung für nicht möglich und/oder nicht sinnvoll gehalten wird, dann bleibt für den der überhaupt über die Richtigkeit von Normen argumentiert, nur die bedingte Begründung, eine Begründung also relativ auf selbst nicht begründete oder begründbare Prämissen, denen zustimmen man andere gerade nicht rational motivieren kann ... (Kuhlmann 1986, p. 207, quoted from Bonaunet 1994, p. 412f.).

What is interesting in this quote is the way he rejects the relative justification, which is the strategy that Tugendhat has argued for. He says that the conditional justification must rest on premises that one cannot "rationally motivate" others with. I will not venture to guess what "rational" means here, but it is in any case false that a relative justification is unable to present motives that it would be possible for all to identify with. Tugendhat has relied on the basic distinction between coercion and equal respect: The arguments in favour of relations of respect refer to the possibility of treating each other as possible subjects of blame or praise and not viewing each other as mere instruments to reach our private ends. This is not an appeal to an intuition or particularistic premises. But whether an unconditional "must" makes better sense than this relative account or not is the issue of Part 2.

## *Part 2*

In Part 1 a grasp of the central points in Tugendhat's moral philosophy was gained and I will now proceed to evaluate some of the criticism this concept of morality has received. Part 2 consists of two chapters; chapter 6 deals with the idea Habermas has of moral an unconditional obligation in terms of epistemic reasons. Chapter 7 discusses Stemmer's attempt to justify a "rationally necessary" obligation while rejecting the standard of universalism. Both authors deny that Tugendhat can connect an emotional account of obligation to the standard of universalism. It will be demonstrated that this criticism relies on implausible accounts of moral obligation and that the arguments against an emotional obligation to universalism are inadequate.

## *6. Sanctions and Obligation*

Habermas has voiced a critique that finds resonance in many readers: Moral obligation is absolute and not relative to our feelings of shame and guilt. The reactive attitudes are derivative of the obligation we have in light of the "validity" of the moral norms. He also denies that the justification of the standard of universalism needs the foundation of reasons in terms of motives. Prudential motivation and moral obligation are completely independent notions, he thinks.

I will look at this in three steps: The first section (6.1) presents the criticism and gives a short sketch of how it fits with Habermas' discourse ethics. 6.2 is an attempt to show that Habermas' theory of obligation is in tension with his own theory of validity, and that his solution of an appeal to a "dogmatic lifeworld" is unrealistic. 6.3 is a further elaboration of Tugendhat's philosophy in order to show that Habermas' critique relies on a misguided interpretation (which I will describe as a "sequential" as opposed to "conceptual" reading).

### *6.1 Dissolution of Obligation?*

Habermas does want to take issue with the central place that Tugendhat grants universalism in morality, he rather doubts that our obligation to universalism can be understood as dependent on prudential evaluations. In Part 1 we saw how Tugendhat explains that the moral "must" depends on at least an implicit decision on how to live. Habermas thinks that this leads to dissolution of the obligation that we have to the moral standards:

"Wenn aber jeder jederzeit aus seiner egozentrischen Perspektive darüber befinden kann, ob es sich für ihn lohnt, sich überhaupt auf Moral einzulassen, verlieren die innerhalb des moralischen Sprachspiels möglichen Züge ihre kategorische Verbindlichkeit. Moralische Urteile und Normen verändern dadurch, daß sie nur relativ zum Ergebnis einer instrumentellen Überlegung Geltung beanspruchen dürfen, ihren illokutionären Sinn." (2004, p. 304n7).

The essence of Habermas' critique is that Tugendhat does not preserve the absolute necessity that we encounter in the "moral language game". The moral restrictions we place each other under are no longer absolute or categorical if the agents stand free to evaluate whether it is in their own interest to belong to the moral practice at all. The

way Habermas sees it moral obligation can only be understood if we presuppose that agents do not weigh their prudential interests against the moral demands. Morality cannot be dependent on instrumental reasoning.

One needs to watch out for the switch in terminology that occurs when discussing Habermas. The quote above relies on his idea that norms and moral judgements have "Geltung" or validity independent of any kind of sanctioning. This is due to his consensus theory of truth and moral "rightness", which allows him to speak of the justification of both assertions and norms in much the same way. "Rightness" is a property that would create consensus in an ideal speech situation, where all epistemic constraints are removed and all possible arguments have been weighed against each other. We will never enter this ideal speech situation, but insofar as we want to argue sincerely we can only use arguments that we think would find acceptance "in the long run". By accepting a moral judgement we also believe that all other rational people would accept it. This is a regulative idea that explains what it means to commit to a judgement.

I will try to answer Habermas' critique independently of any extensive assessment of his special concept of moral validity, to the extent his treatment of the notion of obligation can be separated from this account. As will become apparent, I do not need to be concerned as to whether the truth-analogue concept of "rightness" is plausible or not, but can concentrate on the relation between moral obligation and knowledge in general. The suspicion is that Habermas needs to presuppose a certain level of ignorance on the part of the agents in order to keep the absoluteness of moral imperatives intact. Let us now look at another quote that says something similar to the first quote, but reveals more of the kind of reasons Habermas has to refuse Tugendhat's account:

"Aus begrifflichen Gründe kann der kategorischen Sinn moralischer Verpflichtung nur so lange intakt bleiben, wie den Adressaten die Möglichkeit verwehrt ist, auch nur virtuell jenen Schritt hinter die moralische Gemeinschaft zurückzutreten, der nötig ist, um aus dem Abstand und aus der Perspektive der ersten Person die Vor- und Nachteile einer Mitgliedschaft überhaupt abzuwägen" (1996, p. 35f.).

Here it becomes evident that Habermas thinks that the theorist must make room for an unconditional practical necessity because it belongs to the very concept of moral



obligation (cf. "Aus begrifflichen Gründe"). Tugendhat has a fallacious account of the moral "must" because the *idea* of moral obligation is that it cannot be relative to our personal well-being. Evidently, moral imperatives are hypothetical if they are made out to be conditional upon on the prudential reasoning of the agents. But Habermas says that moral obligation *means* a necessity not relative to our will, and Tugendhat cannot explain this.

This criticism relies on the concept of an illocutionary act. When we present each other with a moral judgement we mean to say that it must be adhered to, no matter what preferences the agent has. The idea is that it would not be meaningful to perform a speech act that claimed the absolute necessity of some moral standard, while at the same time considering this standard to be relative to the prudential reasoning of the addressee. Both the speaker and the hearer need to be involved in the speech game in such a way that they do not weigh the moral imperatives against private interests. Moral obligation would dissolve if it were to be understood in the quasi-contractarian way that Tugendhat has explained it since it is the nature of moral imperatives is to be unconditional.

The claim that Tugendhat does not explain how moral imperatives are unconditional to the addressee is correct, of course. He does not think there is such a thing as an absolute practical necessity. But does this mean that he cannot speak of moral obligation either? If by moral obligation one understands unconditional necessity, then yes. And the illocutionary acts of our moral language game seemingly provide a reason to understand obligation in this sense. Looking at the last quote again, one can see how Habermas claims that a philosophical investigation needs to leave the idea of an unconditional obligation *intact* because it is the precondition of meaningful moral interaction. But is this the right place to start?

The goal of having a theory of the moral practice suited to a literal understanding of illocutionary acts is a dubious affair. That would be the wrong "direction of fit", so to speak: A philosophical illumination cannot be judged by whether the resulting understanding of the moral practice fits the grammatical absoluteness of our imperatives. Instead, the proper starting point involves a comprehension of our seemingly "absolute imperatives" in light of a general theory of human practice. Language game must fit theory of action, not the other way around, since our moral speech games may refer to a practical necessity that cannot be made

philosophical sense of. We cannot take our linguistic expressions at face value, and we cannot wrestle an account of moral obligation out of the mere grammatical form of moral imperatives. The fact that we do not usually mention the consequences of transgression when we present each other with moral demands does not mean these are something secondary.

So Habermas needs to present an alternative account of the moral "must" that demonstrates how sanctions can be said to be non-essential to our moral obligation. According to Tugendhat, there has only been one alternative account of the moral "must" and it was Kant's (Tugendhat 2001, p. 8). But in order to latch onto this account one needs a "transcendental" concept of reason and Habermas does not want to carry the load of this kind of metaphysics (cf. Habermas 2004, p. 302). But he still wants to preserve Kant's interpretation of moral obligation as "Achtung vor dem Gesetz" (cf. *ibid.*, p. 304). He wants to claim that it is knowledge of moral reasons that obligates us, not our sense of shame. He therefore speaks of "Sollgeltung", a normative force that moral norms have in virtue of being the possible objects of universal consensus. But this needs to be explained without presupposing that we have a higher, "transcendental" nature. In the following I will give a very short description of how Habermas envisions such an account.

Habermas does not deny that the Strawsonian Triad is important, but claims that these reactive attitudes are secondary to the notion of obligation and not constitutive, like Tugendhat thinks. They should be understood as "Warnsignale" (cf. Habermas 2004, p. 306), meaning that they are intuitive evidence that some obligation has been violated. How is this obligation constituted independent of our emotions, then? The unconditional nature of moral norms arises from the way justifiable moral reasons govern us: "Gründe sind aus einem besonderen Stoff; sie zwingen uns, mit Ja oder Nein Stellung zu nehmen. Damit ist in die Bedingungen verständigungsorientierten Handelns ein Moment Unbedingtheit eingebaut" (Habermas 1984, p. 27).

Not sanctions, then, but the fact that we are "always already" participating in a communicatively structured lifeworld of reasons constitutes the moral "must": "Die Fortsetzung kommunikativen Handelns mit diskursiven Mitteln gehört zur kommunikativen Lebensform, in der wir uns alternativlos vorfinden" (*ibid.*, p. 345). The standards of a rational discourse pervade our lifeworld - they are a part of our

social fabric. And importantly; the norms that have emerged from discourse are taken to be certain as the knowledge of the empirical reality (cf. *ibid.*). It is assumed that in order to get by in everyday life agents must naively consider their mere opinions to be absolute truths. Just as we need to forget about the fallible nature of all empirical propositions in order to dare get into an elevator, we must also take moral judgements to be unconditionally right.

So agents involved in the moral language game do not act morally to avoid shame and guilt. On the contrary, all that is needed is the "zwanglosen Zwangs guter Gründe, mit dem sich moralische Einsichten als Überzeugungen einprägen" (Habermas 1991, p. 151). "Achtung vor dem Gesetz" now means the ability to be influenced by the outcome of a rationale discourse. Obligation is explained without recourse to our self-conception and the way morality is connected to our pre-moral interests. The moral discourse is fully contained on the level of reasons that aspire to universal validity.

Habermas says this is necessary in order to preserve the *cognitive* nature of moral judgements, where cognitive means aspiring to reach consensus in an ideal speech situation. If the norms were thought of as relative to our self-conception they would no longer be moral but "ethical", a term Bernard Williams used to describe the kind of questions that concern how to live an authentic and fulfilling life (cf. Williams 1985). The norms would thereby belong to a doctrine of the good life, which people may or may not share. Habermas warns against letting morality be perceived as a club that fits only those with a specific idea of the good life. Moral norms can only be cognitively "valid" if they do not rely on any particular "ethical" ideas of a fulfilling existence.

But isn't this cognitive purification of moral obligation an attempt to fit the world to a mere idea of how one would like it to be? This suspicion could be avoided if it could be empirically verified that the world matches this understanding of moral obligation. And Habermas claims to have such verification: "Psychologische Untersuchungen belegen, daß Kinder schon früh lernen, unbedingte moralische Verbote von anderen sozialen Regeln und bloßen Konventionen zu unterscheiden" (Habermas 2004, p. 306). This is supposed to demonstrate that people actually perceive moral norms to be absolute restrictions. Children know that what moral validity is of some radically different nature than mere conventions, since the

rightness of norms is unconditional. There is no need to squeeze a notion of sanctions in between.

Now the contours of what seems to be a third account of our moral obligation have been indicated. It neither relies on Kant's concept of a transcendental reason nor the "inner sanctions" of self-assessment. It places its trust in the idea that we can be governed by reasons that could find a consensus in an "ideal speech situation". Our obligation is to the rationality of the moral norms; we do not need to add anything to the perception of epistemic reasons in order to identify the nature of our obligation. This explanation of the normativity of universalism claims to preserve a pre-theoretic notion of moral behaviour being unconditionally right. If this pre-theoretic understanding also makes theoretical sense, then Tugendhat's account of the moral "must" has conceded too much to pre-moral reasoning.

### *6.2 Theorist and Agent*

This third account of obligation in terms of communicative reason must now be evaluated on its own terms. That means that one cannot simply ask: "Wäre die Irrationalität in diesem Sinn eine Sanktion?" (Tugendhat 1993, p. 45) This question presupposes that sanctions are necessary in order to understand moral obligation at all. But as has just been shown, Habermas thinks it isn't. In this section I will discuss the idea of "Sollgeltung" on discourse ethics' own terms, not questioning the idea that moral knowledge alone can be the source of obligation. It will be demonstrated that the idea of an unconditional obligation is implausible even if one allows this idea.

In the section on modernity (1.2) I investigated Tugendhat's notion of "us". I found that the way he left it unspecified was unfortunate. But from a methodological standpoint his use of first person plural is important, at least the first person bit. It signifies that he is standing in the middle of the moral practice that he is discussing. He is not speaking from some extramundane standpoint, but starting from the disorientation that is experienced in the practice itself. He does not speak of the "participants" in the moral practice or the different "lifeworlds" they inhabit, but only the way *we* relate to moral norms.

Habermas, on the other hand, deliberately speaks from the third person perspective. He is the analyser of discourses, not the participant. If it was fitting to put Tugendhat in the shoes of Ivan from *The Brothers Karamazov* (cf. 1.2), then it is just as appropriate to liken Habermas to the role of Dostoyevsky. It is Habermas' thesis

that the reflections that he goes through are incompatible with the constitutive standards of the lifeworld he is analysing. Only "die institutionalisierte Wissenschaft" can neutralise the "Naivität", "Platonismus" and "dogmatische Verfassung der Lebenswelt" (Habermas 2004, p. 320). Why does Habermas use this kind of language to describe the social sphere of interaction?

That is because his concept of an unconditional obligation stands in tension with his own consensus theory of validity. This concept of truth and moral rightness sees validity as a regulative idea, something that can never be finally achieved but must be aimed at in every sincere argument. That means that we can never be sure that we have moral norms that are fully justified. But Habermas needs agents that take the moral norms to be unconditionally valid and therefore obligatory. We are said to take our opinions to be absolutely correct, even though we are fallible beings. His concept of a categorical obligation therefore depends on agents "die nur *zeitweise* die reflexive Einstellung von Argumentationsteilnehmern einnehmen" (ibid., emphasis in original).

His idea of the unconditional moral language game is incompatible with agents that have reflectively incorporated the insights of theory into their practice. Agents may perceive the provisional nature of their norms only in rare moments of moral discourse, and must then submerge themselves back into the unenlightened lifeworld. Here they forget about the fact that this norm can only be finally validated in an ideal speech situation. Habermas does not allow agents to internalise knowledge of the actual workings of the moral sphere. They are defined as necessarily naïve performers.

The underlying idea is that the phenomenologically absolute moral "must" has to be preserved in a philosophical account. The moral language game is conceived as a bubble that would burst if certain considerations were allowed: "Die skeptische Option eines Ausstiegs aus dem Sprachspiel *begründeter* moralischer Erwartungen, Verurteilungen und Selbstvorwürfe besteht nur in der philosophische Reflexion, aber nicht in der Praxis" (ibid, p. 345). But Habermas is putting the theorist in an unwarranted epistemic situation by claiming certain options to be exclusive to philosophy. By making it conceptually impossible for a moral agent to incorporate the possibility of scepticism into the sphere of decision-making he is underestimating the reflective nature of moral interaction.

When introducing the discussion of Tugendhat's account of moral obligation I claimed that his way of finding a formal concept of morality through a discussion of our moral semantics is in accordance with Giddens' methodological criterion of "double hermeneutics" (cf. 2.1). Here I concentrated on the possibility of theoretically interpreting social action. But Giddens had another point in mind as well, and that is the fact that interpretation works both ways: "The concepts and theories invented by social scientists ... circulate in and out of the social world they are coined to analyse" (cf. 1987, p. 19). The agents incorporate the insights of theoretical reflexion into their practice, and in fact their doing so is a strong indication that the hypothesis is supported by reality.

Again, philosophy is not to be equated with social science, but the point Giddens is getting at has philosophical relevance. Habermas could be right - as an empirical fact - that the moral sphere is perceived by agents as governed by absolute notions that are incompatible with a philosophical reconstruction in terms of self-interest. The problem is how he uses such a given understanding of moral obligation to disqualify accounts that do not preserve this intuitive perception. He is saying that since certain philosophical reflections would destroy a pre-reflective conception of moral obligation they are disqualified as possible ways for the agents themselves to understand this concept.

Habermas is right to insist that a philosophical reconstruction must take account of the agent's own perspective. But he understands this the wrong way; he wants the reconstruction to *limit* itself in order to comply with the "dogmatism of the lifeworld". He thinks that the phenomenological basis of morality must be kept intact, and therefore defines practical considerations that clash with this basis as belonging to some whole other discourse. His appeal to illocutionary meaning is the clearest example of this, taking ordinary language at face value. The fact that we have certain categorical ways of speaking is taken to reveal the possible attitudes and reflections.

In the end, this relies on the idea that there is an epistemic distance between theorist and agent. The agents have to be "protected" from the possibility that the theorist sees, that is, deliberating whether it is beneficial to join the moral community at all. Remember that he pointed out how investigations in psychology were said to prove that children learn from early on that moral obligation is absolute (cf. Habermas 2004, p. 306). It is very telling that Habermas decides to support his account of an

unconditional obligation by referring to those who are most likely to have only a *partial* understanding of the moral practice.

It is clear that this cannot satisfy the methodological point of a "double hermeneutics", since this standard is incompatible with such a perception of distance between agent and theorist. Giddens says: "Now every competent social actor is himself a social theorist, who as a matter of routine makes interpretations of his own conduct, and of the intentions, reasons and motives of others as integral to the production of social life" (1976, p. 153). The point is that we must assume that the agents could have the same theoretical framework to their disposition as the philosopher. Or better yet, we should make no distinction between the agent and the philosopher at all.

Perhaps Habermas can allow this in his theory, but he cannot take account of it with the full force it deserves. As we saw earlier, his agents can "*nur zeitweise*" respond to the knowledge of the non-absolute status of moral judgements. There are no good grounds, however, to assume that the fallibility of our opinions cannot be reflexively incorporated into the sphere of interaction. On the contrary, there are strong reasons to allow agents the same relation to the status of our knowledge as "*die institutionalisierte Wissenschaft*".<sup>12</sup> Placing restrictions on the self-knowledge that the agents could have implies that the social concept is not explained coherently. It reveals that the theorist is clutching to an idea that cannot be given a proper theoretical elucidation. In this case it is the idea of an unconditional moral obligation.

In this section it has been shown that Habermas' idea of obligation stands in tension with his own theory of validity, and that he solved this by creating a gap between theorist and agent. If he had decided to close this gap he would not have had a theory of an unconditional practical necessity. This makes it clear that Habermas cannot explain an unconditional "must" even if one grants him the strange idea that obligation consists solely in perception of reasons that stand the test of discourse.

---

<sup>12</sup> Karl-Otto Apel has a similar critique of what he calls Habermas' "*geschichtsabstraktiven* Verständnis des Verhältnisses von präreflexiven Handlungswissen und Diskurs" (2003, p. 195). Although his argument is limited to scientific progress towards truth, it is parallel to what is said above. He argues that Habermas does not take sufficient account of the ability to *internalise* the fallible nature of our judgements and to consider each consensual agreement to be provisional.

### 6.3 *Conceptual Priority, Not Sequential*

The discussion of Habermas has shown that his theory relies on a problematic account of the relation between agent and knowledge. But some of the critique of Tugendhat may still be right even if the idea of an absolute "Sollgeltung" proved to be unsound. The claim was that the moral "must" would be dissolved if it were understood as relative to our prudential considerations on whether to belong to morality at all. The question is how one can understand the practice of giving mutual imperatives when each member can always decide for himself whether to care about the demands he is faced with. The necessity referred to in a speaker's demand of moral conform behaviour does not seem to have a genuine correlate in the addressee.

When put this way the matter seems tricky for Tugendhat. But one important reason for discussing Habermas' criticism is exactly that it welcomes an attempt to make it clear why one should *not* put things this way. The accusations rely on a reading of Tugendhat that views prudential interests as *continually* threatening to disband moral obligation, but he does not see that Tugendhat differentiates the prudential and the moral sphere of reasons as an analytical device that demonstrates conceptual priority. In this section I will seek to clarify what this means.

The argument presented in 5.3 was that the agents must view the standard universalism as a prudential good in some way in order for them to be governed by it. There has to be some volitional identification with the goal of achieving symmetric relations with others, so the strong must view morality as the key to the good life on some level. The leap from tyrant to moral agent may rightly be called "ethical" in the way described above. It regards our perception of how to live; relations of mutual respect are seen as part of what it means to have a rewarding existence. We want to adhere to the standard of universalism in order to keep our self-conception intact, in order not to feel shame. Most people never have to take this step from tyrant to moral agent. But it is clear that some implicit stand on how to live has been taken by all who adhere to the moral standard of universalism. It must be embedded in their view of the good life if they are to be motivated by it (cf. Tugendhat 2003, p. 72).

But Habermas is not sufficiently mindful of the nature of these considerations in his critique of Tugendhat. Let's remind ourselves of how he understood this relation between the prudential and the moral level in the quote above: "Wenn aber jeder jederzeit aus seiner egozentrischen Perspektive darüber befinden kann, ob es sich für ihn lohnt, sich überhaupt auf Moral einzulassen ..." (Habermas 2004, p. 304n7). I will



call this a sequential reading of Tugendhat. It gives the impression that agents are thought to *first* consider whether it is in their pre-moral interest to be a part of the moral practice, and *then* put themselves under an obligation. This would of course dissolve the moral "must", since every moral imperative would be relative to strategic calculations.

The mistake in saying "jederzeit" is that it creates the impression that our moral sentiments are something that can come up for review at wish. This makes the decisionistic aspect of Tugendhat's theory into a choice that can be remade in every situation, instead of the existential orientation that it really is. Habermas' account covers up the fact that we are involved in a philosophical reflection that does not attempt to retrace the actual mental acts that occur in the actual reciprocal justifications of actions. The anthropological preconditions of moral behaviour do not mirror the conscious perspective of the agent. Phenomenologically it seems that we do the right thing because it is right. But in a philosophical reconstruction of the possibility of moral obligation such talk is empty.

What I mean by "conceptual priority" can be demonstrated by recalling Tugendhat's four levels of freedom of the will (cf. 5.4). It is especially the relation between levels two and three that is important here: The second level was the ability to prioritise our ideas of what is good in order to create a coherent and stable volitional identity. The third level of morality was described as a possible configuration of the second level, where one introduces a condition upon one's actions that they do not disrespect the will of others. The point is that the third level presupposes the second one, so the priority I have been speaking of is priority in sense of being a *precondition*.

It is the relation between the second and third level that is presented in a distorted fashion by Habermas. He imagines that Tugendhat's account amounts to making the moral self-conception continuously conditional upon a separate pre-moral egocentric level. But the idea is not that we can flutter up and down between levels; to have moral self-conception is a special configuration of the second level, it means that we have incorporated moral norms into our prudential priorities. One perceives relations of equal moral respect to be worth pursuing. Our feeling of shame when we transgress is a reaction to the fact that we do not live up to the standards that we have constituted our self by.

Tugendhat's project is a reflection upon a practice that is only implicitly understood in society, it does not presuppose that agents have any explicit knowledge of the conceptual interconnections that constitute it or what the motivational basis is. But as we saw in the previous section, one cannot disallow this possibility either. Perhaps the "ethical" embedment of universalism is slowly becoming everyday knowledge. According to Giddens, this is bound to happen if the thesis has purchase on reality. This does not have to lead to dissolution of obligation. Could it not rather make us identify stronger with the standard?

The project is to understand our moral practice (cf. 1.1), so whether the moral "must" turns out to be weak or strong is no measure of adequacy of the enterprise as such. Tugendhat wants to clarify what the various practical options entail, making it evident to us what our decisions mean. Habermas' claim that this will destroy the idea of moral obligation only reveals distrust in our will to incorporate moral norms into our self-conception.

## *7. Contractarianism vs. Wishful Thinking?*

Tugendhat's focus on the way morality is dependent on our prudential interests brings him close to the contractarianism that stands in the Hobbesian tradition. But the strict contractarian account denies that the standard of universalism can be justified without some dubious metaphysical premise. Peter Stemmer faces Tugendhat with a critique that comes from the direct opposite direction of the one put forward by Habermas. Both critics claim that Tugendhat is stuck in an impossible mid-position, and that he has to go with communicative reason all the way or embrace the premise of self-interest to the full. But the critique from the contractarian camp presents a very different kind of task; instead of rejecting strong deontological claims Tugendhat has to add a principle of universalism to the comprehensible and minimal basis of self-interest. And contractarianism does not care much for additions.

### *7.1 Universalism as a "Subjective Ideal"*

Stemmer does not question Tugendhat's account of moral obligation as relative to our will and the "inner sanctions" of conscience. Indeed, his philosophy does to a great extent consist in acknowledging the main tenets of Tugendhat's account of obligation, while demonstrating that universalism does not belong to this concept of morality. He doubts this standard is more than a religious residue:

"Die überkommene, jüdisch-christlich geprägte universalistisch-egalitäre Moral, die allen Menschen einen moralischen Status zuerkennt und von einer grundlegenden, normativ relevanten Gleichheit aller Menschen ausgeht, stützt sich auf metaphysische bzw. religiöse Annahmen, von denen sich nicht zeigen läßt, daß es rational ist, sie für wahr zu halten." (Stemmer 2000, p. 248)

The argument is that universalism presupposes a faith that we cannot prove to be rational. Stemmer asks us to consider the rational sceptic who does not care about the interests of others, does not believe in any moral standards and has no religious beliefs (Stemmer 2000, p. 17). If the moral norms are to have a rational justification they must be justifiable to this creature.

The norms that a rational sceptic would not accept are irrational and reliant upon "subjective ideals". Saying that a norm is dependent on a "subjective ideal" has the same function as when Tugendhat says that a certain justification is dependent on a

belief in a "higher authority". So Stemmer's claim is that it is impossible to justify the standard of universalism without appealing to something beyond our ordinary interests. This is because the rational sceptic accepts what is necessary in order to protect his interests, and only this. If his interests are taken better care of within a moral practice then he has a reason to enter this. Tugendhat's criterion of justification was also that the norms must be in the interest of the agents, but he added the criterion that they have to be in the *equal interest of all*. The argument was that those that are disfavoured by the norms are to some extent subject to coercion, and a coerced norm is the opposite of a justified norm.

But Stemmer would say that all this talk of not subjecting others to force is begging the question. He can grant that universalism is the real alternative to coercion, but forcing others is only bad from a perspective that is already laden with moral ideals. The rational sceptic admits the need for norms that secure that others do not harm him, but he does not believe that norms that stretch beyond this criterion can be justified. It is simply not in his interest to put himself under an obligation towards the harmless. Remember the quote from 5.2: "Wenn A nicht in der Lage ist, B zu verletzen, dann hat B keinen Grund, sich vor der Verletzung durch A dadurch zu schützen, daß er eine Verpflichtung übernimmt, A nicht zu verletzen" (Stemmer 2000, p. 210). It is irrational to introduce a norm of equal respect towards those that couldn't hurt us in the first place.

So it may seem like Stemmer willing to bite a bullet that Tugendhat has dodged in an illegitimate way. Is it the case that contractarianism has faced the bitter consequences of the premises that Tugendhat has started off with? The initial premises seem very similar: It is a matter justifying norms to someone by giving motives. And the two philosophers rely on the same straightforward concept of practical reason and interests. Stemmer has forcefully argued that it is impossible to understand universalism as a rationally necessary criterion on our moral norms, and now the burden of proof seems to lie on Tugendhat.

### *7.2 Who wants Justification?*

In order to see how Stemmer and Tugendhat arrived at such different conclusions we have to look closer at their initial premises. Let's first remind ourselves of why one needs to justify moral norms at all. Remember how Tugendhat used the example of a

child that asks its parents why they get upset (cf. 3.1). The child did not *understand* why such a negative reaction was appropriate. As we have seen in the previous section, the contractarianism that Stemmer advocates does not start with a child but rather with a rational sceptic. How do these different starting points affect the results?

Tugendhat does not tell us anything about the child, except that it does not understand why one should identify with the norm. The child does not have to be convinced that it is rationally necessary to be moral, but it has to recognise that there are good reasons. It is just a placeholder that asks "why?" until the spade has turned, and we do not know anything further about its motivational dispositions. The rational sceptic, on the other hand, is a rigorously designed character. We know that he will not accept any practical "must" that is not rationally necessary. That means that the norms must be justified by reference to "basalen Interessen ... die man jedem Menschen fraglos unterstellen kann" (ibid., p. 209). So the moral "must" is a dictate of reason, in that it would be irrational to prefer a society that did not secure our interests in the way that contractarianism suggests.

The difference can be illustrated by Kant's two concepts of hypothetical imperatives. Kant differentiated between *assertoric* and *problematic* hypothetical imperatives (cf. Kant 1785, p. 414ff.). The assertoric ones refer to an interest that we know is present in all rational beings, e.g. not to be used as a mere means to an end or suffer injury. The problematic imperatives, on the other hand, refer to a *possible* interest, that is, interests that agents may have but that we cannot presuppose them to have. Evidently, if you base your moral norms only in interests that you can presuppose everyone to have it will not be conditional upon how people happen to conceive of themselves. Stemmer's attempt to deduce only the moral norms that a rational sceptic could accept aims at this kind of *rational necessity*. This creates a morality of assertoric hypothetical imperatives. It has the downside that it cannot capture some central moral ideas, such as equality and justice. There is no way to show that it is rationally necessary to treat everyone with equal moral respect or distribute goods according to some morals standard.

Tugendhat does not aim at this strong sort of practical necessity. He is in the business of justifying problematic hypothetical imperatives instead. He wants to demonstrate that there are good reasons to promote an egalitarian concept of morality and to envision an unrestricted moral community. But he has no arguments that can

prove this to be rationally necessary. All that can be done is show what a rejection of universalism implies and how this affects the possibility of interaction as we know it, thereby refuting the idea that they are mere metaphysical remnants of pre-enlightened times. Compared to Stemmer's contractarianism this has the downside that it is possible to find fully rational human beings that do not have an interest in norms based on this standard.

To claim that Tugendhat's justification is problematic as opposed to assertoric is not a fancy way of saying "anything goes". It does not mean that one could in principle justify any kind of moral standard. He has a specific idea of how the moral practice must be understood and why must involve universalism, and given the nature of this social practice it can only be given a justification relative to an interest that we cannot presuppose in all rational beings. Stemmer, on the other hand, has decided in advance that morality is to be justified by assertoric hypothetical imperatives. How can he simply decide this?

He does not give us any explicit reasons but merely assumes that it is only the rational sceptic that demands justification of the norms. This move is too quick, however, and hardly the innocent precondition he makes it appear to be. How do we know that moral norms are the sort of thing that can be justified to a rational sceptic at all? Morality must have some connection to our prudential interests, but contractarianism seems to jump straight from this observation to the claim that morality has to be fully justified on egoistic principles. Stemmer's critique of Tugendhat only seems plausible once this premise is accepted. This leads us to the question as to why one would give credence to it.

The contractarian answer seems to lie in the fact that we have created the moral norms ourselves, and it is up to us to make sure that they are rational enough for us to keep them. The world contains no moral standards independent of us humans; they are devised to protect our interests in the first place. If the moral practice contains standards that are not rationally necessary we ought to consider them unjustified. This means that Stemmer advocates a revisionist view on the moral practice, discarding every feature that cannot be justified when subjected to the critical questions of a rational sceptic.

This leads him to the claim that the sole reason to speak of objective moral standards is a need to suppress their subjective nature: "Man projiziert sein Ideal auf

diese Weise in die Welt und verankert es in etwas objektiv Vorgegebenem" (Stemmer 2000, p. 305). The suspicion is that the norms have been hypostasised in order to have greater force. So Stemmer is in agreement with Nietzsche in saying that this has been a cunning device that helps the weak against the strong.

We must distinguish between Stemmer's claim that there are no objective moral standards and a claim of moral relativism. To say that there is no standard over and above strength and interests is not the same as claiming that the different kinds of moral practices around the globe are equally rational. Stemmer claims that he presents the *one* rational morality, and practices that conflict with this are based on irrational beliefs (cf. 2000, p. 201ff.). So neither Stemmer nor Tugendhat are relativists, their dispute regards whether an obligation to universalism is possible without appealing to a "subjective ideal" or "higher authority". The contractarian claim is that the objectivity that our moral judgements aspire to is a wishful projection.

But one must be careful to take note of the way Stemmer speaks of objective moral standards. The claim that objectivity is a projection of desires into a seemingly worldly existence is a very uncharitable reading of the term. Time and again he speaks of objectivity terms of physical reality. Moral objectivity is denied because there is no ontological realm between personal desires and brute facts. That he perceives of things in this way is revealed by the way he defines normativity: "Normativität entsteht durch das Zusammenkommen zweier Bausteine, eines Müssens der notwendigen Bedingung und eines Wollens" (Stemmer 2008, p. 42). The world of normativity consists of a desire and the necessary conditions to satisfy this desire. Normativity is the pressure that urges us to take the necessary means when we want to reach certain goals. It arises out of the confrontation between the subjective volition and the objective world.

If such a binary scheme is the only sound foundation of morality it is evident that the norms have to be rational in a strictly prudential sense. Since normativity is the combination of a desire and the necessary means to its fulfilment there can be no normative standards that conflict with straightforward instrumental reasoning aimed at what is in the optimal self-interest. Whatever objectivity there is to speak of has to come in the shape of a necessary condition to the satisfaction of a desire. This toolbox does not give us any reason to think that universalism is more than wishful thinking. So the question is whether one has to be willing to accept this stripped-down and

egoistic version of moral obligation in order to justify moral norms. Can Tugendhat both maintain his concept of what morality is, which implies an objective standard, *and* consider it to be justifiable without presupposing some "subjective ideal"?

### *7.3 Impoverished Topography*

Let's first deal with the notion of 'objectivity'. Stemmer claims to be grounding moral norms in the only objective foundation there is: interests and strength. There is no place for universalism within this framework since this standard refers to the option of putting strength aside and dealing with each other as equals. Contractarianism regards the possibility of treating all with moral respect as a contingent ideal, which can only be causally explained and not justified. The causal origin is assumed to be due to some unwarranted leap from equal descriptive features as species to normative equality (Stemmer 2000. p. 298f.).

But this way of clutching to the tangible is blind to the fact that morality is a *collective* enterprise. The standard of universalism loses its air of subjectivity once its place within the structure of the moral practice has been identified. Tugendhat has demonstrated how symmetry gets transformed into a normative demand in virtue of being the opposite of coercion. The opposition between force and symmetry is not between something real and something ideal; both alternatives can be described neutrally and independent of any moral norms. In order to opt for universalism there is no need to project some subjective ideal unto the world. The normative pressure to treat each other on terms of symmetry arises from the nature of our interaction. Nobody wants to be treated as a mere means to an end, nobody wants to be subject to coercion. What is it in Stemmer's theory that blocks this understanding of universalism?

It is his impoverished topography: His map of the moral sphere reveals nothing of the more complex details of human behaviour that must be taken account of when dealing with social instead of individual actions. The reason he takes the moral standard of universalism to be an irrational residue from our religious past is that he uses the individual rational sceptic as a looking glass towards the moral practice. But symmetry is a structure of collective actions, and cannot even be comprehended unless one is willing to move beyond an individualistic understanding of the rationale behind moral norms.



So the notion of 'objectivity' becomes less problematic once the normative standard of universalism is understood as a consequence of the possibility of symmetry in collective actions. As a real alternative in social interaction it would be misleading to call it a "subjective ideal". It has not emerged from some warm feeling of brotherhood, but is perceived as the basic alternative between coercion and agreement. The interesting question is not the extent to which universalism is objective, then. The true conflict between Stemmer and Tugendhat is whether or not the standard can be justified to the person who is not yet a member of the moral practice.

We have already seen that there are good reasons to enter the moral practice based on universalism if one wants to participate in the most basic forms of social interaction. But this practice cannot be justified to the sceptic, because it involves the same moral respect for all regardless of whether this respect would benefit our interests or not. The rational sceptic will only enter a limited moral community that protects certain needs. But one has to ask; what is the point of construing a moral theory that fits the rationality of a sceptic?

The idea behind this contractarian construction is the necessity referred to in the moral "must": If this necessity is to be unconditional it has to refer to some premise that we accept in light of the kind of beings that we are. In this regard Stemmer's project is very similar to Habermas': They both think that it is necessary to understand the moral "must" in an absolute sense. They interpret the categorical nature of our moral vocabulary to mean that our norms must be rational to such a degree that nobody would find it more rational to stay immoral. Even though they have very different ideas of rationality the aim of their theories is the same, in that they both think "daß das moralische Müssen ein rationales Müssen sein muß" (Stemmer 2000, p. 67).

This means that we have to subject Stemmer's account of moral obligation to critical scrutiny as well. If he succeeds in explaining moral obligation as rationally necessary this could be an indication that Tugendhat's understanding of the moral "must" is incoherent. Perhaps it was unwarranted of Tugendhat to think of the normativity referred in moral language as necessarily relative to a will to be moral at all. Maybe the moral "must" refers to a stronger kind of necessity than he thought

possible, one that can be justified on premises that are independent on how we want to live.

This would mean a refutation of the idea discussed in chapter 5.3, which can be summed up in the following quote: "Wir können uns nur auf das Wie einer Lebensweise berufen" (Tugendhat 1999a, p. 182). The claim was that morality could only be justified to the strong by appeal to the qualities of a life in relations of equal respect. If Stemmer has a more plausible account of obligation this implies that moral norms can have a rationally necessary justification. But this strong necessity knocks universalism out of the picture, since such a moral standard depends on a weaker kind of obligation. So the alternative is between universalism and an absolute "must".

#### *7.4 Emotional Substitutes?*

In 2.1 it was shown that the moral sentiments that constitute our obligation have a propositional content that refers to a moral standard. There has to be some shared principle that makes these sentiments more than spontaneous feelings of disapproval. To resent somebody one needs to judge some behaviour to be *bad*. But contractarianism only knows the prudential '*bad-for-me*'. Such considerations are meant to show that the moral "must" cannot be established on a mere contractarian foundation, and universalism was then shown to be the only justifiable standard on modern premises. How does this reflect on Stemmer's account?

As has been mentioned, he has much the same account of moral obligation as Tugendhat, in that he sees it as relative to our interests and to the sanctions. The question is whether he can use the idea of "inner" sanctions and still claim that norms are the mere result of a balance of forces? Is it possible to conceptualise moral obligation in terms of emotions in a way that does not burst the frame of his theory? Stemmer thinks so.

He agrees with Tugendhat that in order for the Strawsonian Triad to be applicable one needs to admit some objective standard that states that certain type of behaviour is immoral, and that this is completely incompatible with the strict contractarianism that he defends (Stemmer 2000, p. 136ff.). But he thinks that the moral "must" can be maintained by substituting resentment with moral *anger* (cf. *ibid.*, p. 139). We cannot claim that transgressor has done something that is morally

bad, but he has done something that shakes the foundation of the norms and that makes the rest of the community angry.

But this fails to provide us with an emotional correlate on the part of the transgressor. It cannot be guilt, which is the correlate to resentment. Perhaps Stemmer thinks of fear, which does seem like the proper complement to anger. But if fear is what obligates us to morality the norms seem to be something external to the agents. The moral obligation is relying on basically the same mechanisms that a street robbery operates on: "Do what I say or you will face the consequences". This is not obligation but extortion, and Stemmer does seem to want to fuse these two notions: "Die Funktion des Pflichtbegriff ist gerade die Abgrenzung gegenüber dem Erpresserischen" (2002, p. 42). A person who hands over his wallet to an armed robber does so out of fear, he does not do it because he is obligated. He would only act out of an obligation if he thought that the robber had a *right* to his money. The robber would acquire this right if the victim could agree to a social arrangement where he could be robbed on the street. We *give* each other moral rights (and thereby take on duties). But how are things changed by bringing rights into the picture?

The fact that the community has a right to be angry does not change the fact that we are missing an emotional correlate on the part of the transgressor. Why should the agent care whether the others are angry with him or not? Stemmer claims the perceived or anticipated anger of others will result in "ein Gefühl des Unbehagens" (2004, p. 150). So it is not fear, but a feeling of unease that constitutes the moral "must". But Stemmer does not give us sufficient reason as to why unease should occur in an agent that has no moral self-conception. Why should the radical sceptic feel discomfort by the thought that others would not like him anymore?

This unease by the thought of the reaction of others presupposes a social identity and a need to be acknowledged as a good person. But such presuppositions belong to a theory like Tugendhat's, and cannot be attributed a rational sceptic. Stemmer tries to avoid this by claiming that a rational person must calculate how continued transgressions are going to affect his life in total, not just in a specific situation. It will lead to social quarantine, and "niemand will das" (2000, p. 159). Is the rational sceptic contained in this "niemand"? The interest in having stable associations with other people should strictly speaking only be attributed to a person

that acknowledges the value of living in relations of mutual respect, that is, relations that built on the standard of universalism.

Perhaps Stemmer can avoid this by saying that association with others is necessary in a mere strategic perspective as well. We all need certain human connections to reach our ends. Be that as it may, the need to anchor moral obligation in such long-term considerations presupposes a relatively static social surrounding. The transgressor must be repeatedly faced with the same group of injured people, and his partners in interaction must know each other in order for a social quarantine to be enabled. To have such an interconnected group of social partners is no "rational necessity" in the present age, however, since the swapping of social networks is done with increasing ease. Stemmer seems to implicitly recognise this, and finally mentions the alternative of substituting informal sanctions with formal ones, which means making norms into laws (cf. *ibid.*, p. 160).

The fact that this option presents itself so naturally like the final solution to the problem of sanctions is not coincidental. Hobbesian contractarianism was intended as a theory of political legitimacy, and Stemmer's account unintentionally shows that it must stay on the level of jurisprudence. The contractual idea of legitimacy in terms of interest cannot be customized to an understanding of the moral sphere without sewing notions of identification with an objective standard into it. The concept of universalism as an objective standard is not optional embroidery on an already sufficient basis; it is essential to the concept of obligation. The affective and intersubjective nature of moral norms cannot be made of the poor fabric of strength and interests. Tailoring metaphors aside; it is now clear that Tugendhat's addition of universalism is not due to some religious yearning, but a structural necessity in order to understand a secular moral "must" at all.

Ironically, it is the moral practice based on universalism that appears to be the rational choice from a self-interested perspective, and not contractarianism. Stemmer's account defeats its own cause, since abolishing the standard of universalism leaves the agents with no moral content that could be internalised. Everything has been reduced to instrumental rationality, leaving behind the possibility of a substantial principle that one can identify with and have moral sentiments connected to. In reality this means that nobody can expect people who can get away with wrongdoing to be moral. We

would have no reason to trust in the conscience of others, and probably even less reason to want each other's company in the first place.

## *8. Conclusion*

Part 1 of this thesis was a discussion of Tugendhat's concept of morality and what he considers to be the justification of restricting oneself according to the standard of universalism. There were difficulties internal to his theory, but the resources to handle this could be found within his own framework. Obligation to universalism was identified as dependent on a self-conception that is foreign to taking advantage of others.

Part 2 considered the option of an unconditional obligation to morality. Habermas' criticism of Tugendhat was that obligation is not to be understood as relative to our prudential good. The idea of an unconditional obligation in terms of knowledge of reasons that stand the test of discourse was interrogated on its own. I found that the moral "must" cannot be described as unconditional even on Habermas' own premises of moral validity as a regulative idea. His theory of "Sollgeltung" proved to rely on an untenable view on the agent's relation to how the theorist can describe the actual working of the moral practice.

The idea of an obligation without the moral standard of universalism was also rejected. Stemmer's claim that universalism relies on a "subjective ideal" was demonstrated to be the misguided result of an individualistic understanding of morality. The idea of a rationally necessary moral "must" was found to be bereft of the content that could establish "inner sanctions". Replacing the Strawsonian Triad with "moral anger" on part of the offended failed to provide an emotional correlate in the transgressor. In the end it proved to rely on the formal sanctions of law and not the informal mechanisms of morality.

So Tugendhat's view of the relation between obligation and universalism has been demonstrated to be defensible in light of the criticism he received from the two opposing camps. The moral "must" proved to be of a weaker kind than the one imagined in the accounts of both Habermas and Stemmer. But their respective accounts were demonstrated to be incoherent: Neither the idea of "Sollgeltung" nor "moral anger" were unproblematic notions even when the premise of seeing morality as governed by their respective accounts of rationality was granted.

Instead of worrying that the moral "must" is not absolute one should look at the positive outcome of Tugendhat's account. The result of his project is in effect an argument against all claims along the lines that "human solidarity is simply the fortunate happenstance of creation of modern times" (Rorty 1989, p. 68). This view

fails to recognise the anthropological basis of universalism. The option of equal respect for all is the condition of cooperating without coercion. And - when there is no shared belief in a moral authority that could say otherwise - this implies equal moral rights. As we have seen, one does not have to start from "within a certain tradition" as Rawls claimed (cf. 1.2), since the practice that ethnologists identify as morality has certain preconditions that are independent of "our Western culture".

By demonstrating how our intersubjective attitudes are dependent on the standard of universalism Tugendhat has provided a clear account of what is at stake. In order to reject this account of the justification of moral norms one would have to present an alternative account of the moral "must", which implies a different account of the moral sphere itself. This may be possible, but the rationale behind such an endeavour cannot be that we would prefer an account of moral obligation that is stronger than one that makes it out to be relative to our own decision. The impetus to revise Tugendhat's account must come from an alternative perception of how the different features of morality hang together, not restrained by a pre-conceived notion of norms as unconditional.

But sticking to Tugendhat's account has the advantage of not prejudicing ethical ideas: It can incorporate moral concepts like empathy, virtues, happiness and recognition, for instance. Some of this has been touched upon in this thesis, but the ability to include ideas should not cover up the fact that this concept faces difficulties when faced with certain claims of moral phenomenology. There is for instance a tension between Tugendhat's deontological understanding of morality and the thought that morally good people *genuinely care* for each other, that is; they are *compassionate*. It is awkward to say that we have an obligation to have an emotion that is praised when it is spontaneous. This point was voiced in Bernard Williams' famous claim that the idea of an obligation to universalism "provides the agent with one thought too many" (1981, p. 18). Tugendhat's answer to this tension runs the danger of understanding compassion in overly behaviouristic terms, but this discussion cannot be pursued here.<sup>13</sup>

---

<sup>13</sup> Tugendhat's suggestion is that the reciprocal approval of compassion turns it into a moral demand, since the actions that compassion leads to are morally good ones (cf. Tugendhat 2006, p. 28). Those that are compassionate appear as the more genuinely good persons, and so it comes about that we are all expected to expand the circle of people we care about.

Instead, I will end this thesis with a short note on how Tugendhat's concept of morality can incorporate and help us understand the significance of actual developments in society. The past few decades have presented us with new ways of shaping our social sphere, giving us the opportunity to interact with people that are not physically present. The Internet allows for actions that have consequences not shy of what we can achieve when offline. The difference is that we can be anonymous. This resembles the situation of Gyges in Plato's *Republic*, who had a ring that turned him invisible (II, 359f.). Since nobody would react to his behaviour whilst invisible he could bring about all sorts of mischief.

It is doubtful whether an intersubjective sphere robbed of the visible reactive attitudes of others is capable of reproducing the social mechanisms that provide us with a conscience. Hubert Dreyfus, for instance, is adamant that the net cannot make us commit to anything on the existential level that morality demands: "The press and the Internet are the ultimate enemy of unconditional commitment" (2001, p. 88f.). He sees the plethora of chat rooms that one can experiment in as the triumph of nihilism. It is not necessary to follow any standards of cooperation, since social networks can be changed by a couple of clicks. The Internet is a sphere within which any idea can find approval.

But the Internet is surely the perfect arena to pursue the idea of a justification of moral norms on equal terms as well. We can learn about the will and opinions of all, since no spatial boundaries justify our lack of knowledge anymore. The radically egalitarian structure of the net provides an opportunity to come to a common mind in absence of the traditional mechanisms of power (access, knowledge, appearance, strength etc.). It can provide an arena for "a discourse of reason outside power, which nevertheless is normative power" (Taylor 2007, p. 191). This makes electronically mediated spheres of interaction interesting from the perspective of moral theory: They seem to be bereft of the social mechanisms that can maintain the moral "must", but are ideally suited to a process of justification governed by the principle of universalism alone.

So there is a gap between the optimal conditions for an expansion of the process of justification - which is required by the principle of universalism - on the one hand, and the affective basis of moral obligation on the other. This is a problem that can be handled within Tugendhat's theory precisely because of the way he has



conceptualised the way the norms relate to our reactive attitudes. It helps us see how the face-to-face socialisation and egalitarian process of justification are complementary: The moral legitimacy of an impersonal discourse is dependent on a prior affective commitment to consensual relations, which can only result from experiences made in face-to-face interaction. None of the two competing accounts that have been discussed have the resources to perceive this tension.

This shows how this way of seeing relation between universalism and obligation can be rewarding when trying to understand the moral importance of societal developments, as well as when involved the more abstract endeavour to clarify notions such as 'normativity' or 'justice'. Tugendhat's concept of morality is "open", it allows for several further investigations. This versatility must not be confused with vacuity, as we have seen that even two of the accounts that stand this concept closest contain aspects that are deeply irreconcilable with its main ideas.

## Literature

- Apel, K.-O. (1973). *Transformation der Philosophie*, vol. 2. Suhrkamp, Frankfurt a/M.
- (2003) "Wahrheit als regulative Idee" in Böhler/Kettner/Skirbekk (ed.): *Reflexion und Verantwortung. Auseinandersetzungen mit Karl-Otto Apel* (2003, pp. 171-196) Suhrkamp Verlag, Frankfurt a.M.
- Aristotle (1957). *The Politics*. Trans.: T. A. Sinclair & T. J. Saunders. Penguin, Harmondsworth.
- Berlin, I. (1969). "John Stuart Mill and the Ends of Life" in I. Berlin: *Four Essays on Liberty* (pp. 172-06). Oxford University Press, Oxford.
- Birnbacher, D. (2003). *Analytische Einführung in die Ethik*. Walter de Gruyter, Berlin.
- Bonaunet, K. (1993). *Kants etikk og den transcendentalpragmatiske diskursetikken*. Tromsø.
- Dreyfus, H. (2001). *On the Internet*. Routledge, London.
- Gauthier, D. (1986). *Morals by Agreement*. Clarendon Press, Oxford.
- Giddens, A. (1976). *New Rules of Sociological Method*. Hutchinson University Library, London.
- (1987). *Social Theory and Modern Sociology*. Polity Press, Cambridge.
- Habermas, J. (1981). *Theorie des kommunikativen Handelns*, vol. 1. Suhrkamp, Frankfurt a/M.
- (1983). *Moralbewusstsein und kommunikatives Handeln*. Suhrkamp, Frankfurt a/M.
  - (1991). *Erläuterungen zur Diskursethik*. Suhrkamp, Frankfurt a/M.
  - (1996). *Einbeziehung des Anderen*. Suhrkamp, Frankfurt a/M.
  - (2004). *Wahrheit und Rechtfertigung*. Erweiterte Ausgabe. Suhrkamp, Frankfurt a/M.
  - (2008). "Die Dialektik der Säkularisierung" in *Blätter für deutsche und internationale Politik* (4/2008, pp. 33-46). Blätter Verlagsgesellschaft, Berlin.
- Hegel, G.W.F. (1821). *Grundlinien der Philosophie des Rechts*. Suhrkamp, Frankfurt a/M.
- Hume, D. (1888 (1978)). *A Treatise of Human Nature*. L.A. Selby-Bigge (ed.), 2nd edition by P.H. Nidditch. Clarendon Press, Oxford.
- Kant, I. (1785 (1994)). *Grundlegung zur Metaphysik der Sitten*. Felix Meiner Verlag,

Hamburg.

- Kohlberg, L. (1981). *The Philosophy of Moral Development*, vol. 1. Harper & Row, San Francisco.
- Korsgaard, C. (1996). *Creating the Kingdom of Ends*. Cambridge University Press, Cambridge.
- Kuhlmann, W. (1986). "Moralität und Sittlichkeit. Ist die Idee einer letztbegründeten normativen Ethik überhaupt sinnvoll?" in W. Kuhlmann (ed.): *Moralität und Sittlichkeit*. Suhrkamp, Frankfurt a/M.
- Mackie, J.L. (1977). *Ethics. Inventing Right and Wrong*. Penguin Books, Harmondsworth, Middlesex.
- Mill, J.S. (1861 (1972)). *Utilitarianism, On Liberty and Considerations on Representative Government*. Everyman's Library, London.
- Moore, G.E. (1903 (2004)). *Principia Ethica*. Dover Publications, Mineola (NY).
- Nietzsche, F. (1885 (1988)). *Jenseits von Gut und Böse*. Reclam Verlag, Stuttgart.
- Rawls, J. (1971 (1999)). *A Theory of Justice*. Revised Edition. Oxford University Press, Oxford.
- (1985). "Justice as Fairness: Political not Metaphysical" in *Philosophy and Public Affairs* (3 vol. 14, pp. 223-251). Princeton University Press, Princeton.
  - (1993). *Political Liberalism*. Columbia University Press, New York.
  - (1997). "The Idea of Public Reason Revisited" in J. Rawls: *Political Liberalism. Expanded Edition* (pp. 435-490). Columbia University Press, New York.
- Rorty, R. (1989). *Contingency, irony, and solidarity*. Cambridge University Press, Cambridge.
- Scarano, N & Suárez, M. (2006). "Einführung" in Scarano/Suárez (ed.) *Ernst Tugendhats Ethik* (2006, pp. 7-12). C.H. Beck, München.
- Scanlon, T.M. *What We Owe to Each Other*. Belknap Press, Cambridge (MA).
- Smith, A. (1790, (2002)). *The Theory of Moral Sentiments*. 6th edition, edited by Knud Haakonssen. Cambridge University Press, Cambridge.
- Stevenson, C. L. (1937). "The Emotive Meaning of Ethical Terms" in *Mind* (181 vol. 46, pp. 14-31).
- (1944). *Ethics and Language*. Yale University Press, New Haven.
- Stemmer, P. (2000). *Handeln zugunsten anderer*. Walter de Gruyter, Berlin.

- (2002). "Der Begriff der moralischen Pflicht" in A. Leist (ed.): *Moral als Vertrag?* (2002, pp. 37-70). Walter de Gruyter, Berlin.
- (2002). "Moralischer Kontraktualismus" in *Zeitschrift für Philosophische Forschung*, (56 vol 1, pp. 1-21). Klostermann, Frankfurt a/M.
- (2004). "Die Rechtfertigung moralischer Normen" in *Zeitschrift für Philosophische Forschung* (58 vol. 5, pp. 483-504). Klostermann, Frankfurt a/M.
- (2008). *Normativität*. Walter de Gruyter, Berlin.

Taylor, C. (2007). *A Secular Age*. Belknap Press, Harvard.

Taylor, G (1985). *Pride, Shame, and Guilt*. Clarendon Press, Oxford.

Tugendhat, E. (1976). *Vorlesungen zur Einführung in die sprachanalytische Philosophie*. Suhrkamp, Frankfurt a/M

- (1979). *Selbstbewußtsein und Selbstbestimmung*. Suhrkamp, Frankfurt a/M.
- (1981a). "Drei Vorlesungen über Probleme der Ethik" in E. Tugendhat: *Probleme der Ethik* (1984, pp. 57-181). Reclam, Stuttgart.
- (1981b). "Sprache und Ethik" in E. Tugendhat: *Philosophische Aufsätze* (1992, pp.275-314). Suhrkamp, Frankfurt a/M.
- (1985). "Habermas on Communicative Action" in E. Tugendhat: *Philosophische Aufsätze* (1992, pp. 433-440). Suhrkamp, Frankfurt a/M.
- (1989a). "Überlegungen zur Methode der Philosophie aus analytischer Sicht" in E. Tugendhat: *Philosophische Aufsätze* (1992, pp. 261-272). Suhrkamp, Frankfurt a/M.
- (1989b). "Zum Begriff und zur Begründung von Moral" in E. Tugendhat: *Philosophische Aufsätze* (1992, pp. 315-333). Suhrkamp, Frankfurt a/M.
- (1993). *Vorlesungen über Ethik*. Suhrkamp, Frankfurt a/M.
- (1996). "Gibt es eine moderne Moral?" in *Zeitschrift für philosophische Forschung* 50 1/2 (pp. 323-338). Klostermann, Frankfurt a/M.
- (1997a). *Dialog in Leiticia*. Suhrkamp, Frankfurt a/M.
- (1997b). "Egalitarismus und Universalismus" in M. Willascheck (ed.): *Ernst Tugendhat: Moralbegründung und Gerechtigkeit* (pp. 3-25). LIT, Münster.
- (1998). "Was heißt es, moralische Urteile zu begründen?" in E. Tugendhat: *Aufsätze 1992-2000* (2001, pp.163-184). Suhrkamp, Frankfurt a/M.

- (1999a). "Wie sollen wir Moral verstehen?" in E. Tugendhat: *Aufsätze 1992-2000* (2001, pp. 163-184). Suhrkamp, Frankfurt a/M.
- (1999b). "Wir sind nicht fest verdrahtet" in E. Tugendhat: *Aufsätze 1992-2000* (2001, pp. 138-162). Suhrkamp, Frankfurt a/M.
- (2001). "Vorrede". in E. Tugendhat: *Aufsätze 1992-2000* (2001, pp. 7-9). Suhrkamp, Frankfurt a/M.
- (2003). *Egozentrität und Mystikk*. C.H. Beck, München.
- (2006a). "Das Problem einer autonomen Moral" in Scarano/Suárez (ed.) *Ernst Tugendhats Ethik* (2006, pp. 13-30). C.H. Beck, München.
- (2006b). "Erwiderungen" in Scarano/Suárez (ed.) *Ernst Tugendhats Ethik* (2006, pp. 273-312). C.H. Beck, München.
- (2007a). "Anthropologie als 'erste Philosophie'" in E. Tugendhat: *Anthropologie statt Metaphysik* (2007, 34-54). C.H. Beck, München.
- (2007b). "Der Ursprung der Gleichheit in Recht und Moral" in E. Tugendhat: *Anthropologie statt Metaphysik* (2007, pp. 136-155). C.H. Beck, München.
- (2007c). "Willensfreiheit und Determismus" in E. Tugendhat: *Anthropologie statt Metaphysik* (2007, 57-73). C.H. Beck, München.
- (2007d). "Retraktionen zur intellektuellen Redlichkeit" in E. Tugendhat: *Anthropologie statt Metaphysik* (2007, 114-135). C.H. Beck, München.

Williams, B. (1981). *Moral Luck*. Cambridge University Press, Cambridge.

- (1985). *Ethics and the Limits of Philosophy*. Harvard University Press, Harvard.